

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 10:55:58 ; Search time 66.0263 Seconds  
(without alignments)  
76.150 Million cell updates/sec

Title: US-09-766-412-30  
Perfect score: 66  
Sequence: 1 QPVLHLVALNTPFL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	AAB74256	Aab74256 Exemplary
2	66	100.0	13	AAB80862	Aab80862 Angiostat
3	66	100.0	13	ABG97539	Abg97539 Antiangio
4	66	100.0	20	AAB16596	Aab16596 N-termina
5	66	100.0	20	AAY94321	Aay94321 Mouse end
6	66	100.0	20	AAM48821	Aam48821 Endostat
7	66	100.0	20	ADA43206	Ada43206 Mouse end
8	66	100.0	20	ADM99393	Adm99393 Human end
9	66	100.0	23	AAB49806	Aab49806 Murine en
10	66	100.0	39	AAB35584	Aab35584 Antiangio
11	66	100.0	42	AAB71931	Aab71931 Murine si
12	66	100.0	183	AAM49504	Aam49504 Mouse end
13	66	100.0	184	AAY18409	Aay18409 Endostat
14	66	100.0	184	AAY08689	Aay08689 Murine en
15	66	100.0	184	AAY70258	Aay70258 Murine en
16	66	100.0	184	AAB49380	Aab49380 Murine en
17	66	100.0	184	ABG31793	Abg31793 Human end
18	66	100.0	184	ABU64290	Abu64290 Murine en
19	66	100.0	185	AAY06197	Aay06197 Anti-angi
20	66	100.0	191	3 AAB28398	Aab28398 Murine en
21	66	100.0	191	5 AAY7950	Aay7950 Amino aci
22	66	100.0	207	4 AAE02031	Aae02031 Murine en
23	66	100.0	207	4 AAB71930	Aab71930 Murine en
24	66	100.0	207	5 ABB79902	Abb79902 Mouse end
25	66	100.0	207	8 ADN00602	Adn00602 Murine en

26	66	100.0	207	8 ADO43904	Ado43904 Amino aci
27	66	100.0	218	2 AAY08691	Aay08691 Murine ge
28	66	100.0	580	2 AAY08692	Aay08692 Murine ge
29	66	100.0	684	2 AAY25114	Aay25114 Mouse alp
30	66	100.0	1288	2 AAW26328	Aaw26328 Mouse alp
31	66	100.0	1288	2 AAW92297	Aaw92297 Mouse alp
32	66	100.0	1315	8 ABO84585	AbO84585 Mouse can
33	66	100.0	1527	8 ABO84584	AbO84584 Mouse can
34	62	93.9	20	4 AAB49850	Aab49850 Human end
35	62	93.9	20	7 ADK40308	Adk40308 Collagen
36	62	93.9	25	4 AAB49866	Aab49866 Human end
37	62	93.9	44	5 ABG31791	Abg31791 Human end
38	62	93.9	44	5 ABG31787	Abg31787 Human end
39	62	93.9	178	3 AAY94324	Aay94324 Alternate
40	62	93.9	178	4 AAU00900	Aau00900 Human End
41	62	93.9	179	4 AAU00901	Aau00901 Human End
42	62	93.9	180	4 AAU00899	Aau00899 Human End
43	62	93.9	181	4 AAU00898	Aau00898 Human End
44	62	93.9	182	3 AAY59622	Aay59622 Human end
45	62	93.9	182	3 AAY94323	Aay94323 Human end

ALIGNMENTS

RESULT 1  
AAB74256  
ID AAB74256 standard; peptide; 13 AA.  
XX AC AAB74256;  
XX DT 23-MAY-2001 (first entry)  
XX DE Exemplary anti-angiogenic peptide #30.  
XX KW Anti-angiogenesis; plasminogen; VEGF; vascular endothelial growth factor;  
XX LW FLT-1; tumour; metastasis; cancer.  
XX OS Homo sapiens.  
XX PN WO200118030-A2.  
XX PD 15-MAR-2001.  
XX PF 01-SEP-2000; 2000WO-SG000131.  
XX PR 03-SEP-1999; 99SG-00004310.  
XX PA (UYSI-) UNIV SINGAPORE NAT.  
XX PI Gø R, Kini RM;  
XX DR WPI; 2001-257785/26.  
XX PT Peptides comprising a portion of a protein selected from plasminogen, endostatin, VEGF, FLT-1 and KDR/FLK-1 are useful for treating primary tumor growth.  
XX PS Claim 7; Page 18; 34pp; English.  
XX CC The present invention relates to anti-angiogenesis peptides from a portion of a selected from plasminogen, endostatin, VEGF or FLT-1. The invention is used to prevent or treat primary tumour growth or metastasis or undesired angiogenesis  
XX SQ Sequence 13 AA;  
Query Match 100.0%; Score 66; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.5e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPFL 13  
|||||

Db	1 QPVLHLVALNTPL 13	
RESULT 2		
AA80862		
ID	AA80862 standard; peptide; 13 AA.	
XX		
AC	AA80862;	
XX		
DT	29-MAY-2001 (first entry)	
XX		
DE	Angiostatin-derived peptide #19.	
XX		
KW	Cytostatic; angiogenesis inhibitor; tumour metastasis; tumour growth.	
XX		
OS	Unidentified.	
XX		
PN	US6200954-B1.	
XX		
PD	13-MAR-2001.	
XX		
PF	30-AUG-1999; 99US-00385442.	
XX		
PR	04-SEP-1998; 98US-0099313P.	
XX		
PA	(UYSI-) UNIV SINGAPORE NAT.	
XX		
PI	Ge R, Kini RM;	
XX		
DR	WPI; 2001-234520/24.	
XX		
PT	Peptides having potent anti-angiogenic activity, useful for the treatment	
PT	of tumors, comprise a portion of a FLR-1 protein.	
XX		
PS	Disclosure; Col 15-16; 21pp; English.	
XX		
CC	The present sequence is a peptide which is effective in inhibiting	
CC	undesirable angiogenesis. Angiogenesis is the process of new blood vessel	
CC	formation from pre-existing vessels. Inappropriate angiogenesis is	
CC	associated with various pathological conditions including solid tumour	
CC	growth and metastasis. The present peptide can be used to prevent	
CC	tumour metastasis or inhibit the growth of a primary tumour	
XX		
SQ	Sequence 13 AA;	
Query Match	100.0%; Score 66; DB 4; Length 13;	
Best Local Similarity	100.0%; Pred. No. 3.5e-05;	
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QPVLHLVALNTPL 13	
DB	1 QPVLHLVALNTPL 13	
RESULT 3		
ABG97539		
ID	ABG97539 standard; peptide; 13 AA.	
XX		
AC	ABG97539;	
XX		
DT	16-DEC-2002 (first entry)	
XX		
DE	Antiangiogenic peptide, Endo-2.	
XX		
KW	Angiogenesis; plasminogen; endostatin; Endo-; VEGF;	
KW	vascular endothelial growth factor; FLR-1; kinase insert domain; FLK-1;	
KW	angiogenic inhibitor; Angio-; endothelial cell; proliferation;	
KW	tumour growth; blood vessel formation; migration; tubule structure;	
KW	embryonic development; wound healing; tumour metastasis;	
KW	rheumatoid arthritis; psoriasis; anticancer; therapy;	
KW	antiangiogenic therapy; mitogen; tyrosine kinase receptor.	
XX		
OS	Mammalia.	
XX		
XX		
PN	US2002103129-A1.	
XX		
PD	01-AUG-2002.	
XX		
PF	22-JAN-2001; 2001US-00766412.	
XX		
PR	04-SEP-1998; 98US-0099313P.	
PR	30-AUG-1999; 99US-00385442.	
XX		
PA	(GER/) GE R.	
XX		
PA	(KINI/) KINI R M.	
XX		
PI	Ge R, Kini RM;	
XX		
DR	WPI; 2002-749595/81.	
XX		
PT	Novel peptide for preventing/treating undesired angiogenesis, has a	
PT	portion of a protein e.g. plasminogen, endostatin, and potent	
PT	antiangiogenic activity and endothelial cell proliferation inhibition	
PT	activity.	
XX		
PS	Claim 9; Page 10; 24pp; English.	
XX		
CC	The invention discloses a peptide comprising a portion of a protein e.g.	
CC	plasminogen, endostatin (Endo-1-4), vascular endothelial growth factor	
CC	(VEGF), FLR-1 or kinase insert domain containing receptor/FLK-1. These	
CC	peptides are able to act as potent angiogenic inhibitors (Angio-1-5). These	
CC	inhibiting endothelial cell proliferation and retarding tumour growth.	
CC	Angiogenesis is the process of new blood vessel formation from pre-	
CC	existing vessel, involving endothelial cell proliferation, migration and	
CC	assembly into tubule structures. It plays an important role in normal	
CC	physiological functions such as embryonic development and wound healing.	
CC	Inappropriate angiogenesis is also associated with various pathological	
CC	conditions including tumour growth and metastasis, rheumatoid arthritis	
CC	and psoriasis. Anticancer therapy by inhibiting tumour angiogenesis is	
CC	called antiangiogenic therapy. VEGF is an endothelial specific mitogen	
CC	that functions through two high affinity tyrosine kinase receptors, FLR-1	
CC	and FLK-1 (not defined). Protein-protein interactions are crucial to many	
CC	physiological and pharmacological processes and the peptides disclosed	
CC	interfere with these interactions. The peptides are useful for preventing	
CC	or treating undesired angiogenesis and primary tumour growth or	
CC	metastasis. The sequences presented in ABG97510-ABG97559 are examples of	
CC	antiangiogenic peptides	
XX		
SQ	Sequence 13 AA;	
Query Match	100.0%; Score 66; DB 5; Length 13;	
Best Local Similarity	100.0%; Pred. No. 3.5e-05;	
Matches	13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QPVLHLVALNTPL 13	
DB	1 QPVLHLVALNTPL 13	
RESULT 4		
AAW16596		
ID	AAW16596 standard; peptide; 20 AA.	
XX		
AC	AAW16596;	
XX		
DT	28-JAN-1998 (first entry)	
XX		
DE	N-terminal region of a novel endostatin.	
XX		
KW	Endostatin; angiogenesis; cancer; tumour; rheumatoid arthritis;	
KW	psoriasis; ocular; Osler-Webber Syndrome; myocardial; telangiectasia;	
KW	plaque neovascularisation; haemophilic joint; angiofibroma; inhibitor;	
KW	wound granulation; intestinal adhesion; atherosclerosis; scleroderma;	
KW	hypertrophic scar; cat scratch disease; Rochele minalia quintosa;	
KW	Helicobacter pylori ulcer; birth control; collagen alpha type XVIII.	
XX		

```

OS Mus sp.
XX WO9715666-A1.
XX 01-MAY-1997.
XX 23-OCT-1996; 96WO-US016925.
XX 23-OCT-1995; 95US-0005835P.
XX 02-AUG-1996; 96US-0023070P.
XX 17-SEP-1996; 96US-0026263P.
XX 22-OCT-1996; 96US-00740168.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX O'Reilly MS, Folkman MJ;
XX WPI; 1997-259020/23.
XX Isolated endostatin - useful for treating, e.g. angiogenesis dependent
XX cancers.
XX Claim 3; Page 56; 84pp; English.
XX A novel 20kDa endostatin that specifically inhibits endothelial cell
XX proliferation was isolated and is characterised by its preferred N-
XX terminal amino acid sequence. The present sequence represents the first
XX 20 of these amino acids. The N-terminal sequence corresponds to an
XX internal 20 amino acid fragment found in mouse collagen alpha type XVIII
XX satting at amino acid 1105 and ending at amino acid 1124. The N-terminal
XX amino acid sequence of the inhibitor also corresponds to an internal 20
XX amino acid fragment found in human alpha I type XVII starting at amino
XX acid 1132 and ending at amino acid 1151. Endostatin can be isolated from
XX murine hemangioma. The endostatin can be used to treat
XX angiogenesis related diseases, e.g. angiogenesis dependent cancers,
XX benign tumours, rheumatoid arthritis, psoriasis, ocular angiogenesis
XX diseases, Osler-Webber Syndrome, myocardial angiogenesis, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
XX wound granulation, intestinal adhesions, atherosclerosis, scleroderma,
XX hypertrophic scars, cat scratch disease (Rochelle minalia quitoosa) and
XX Helobacter pylori ulcers. The endostatin can also be used to prevent
XX embryo implantation, i.e. in birth control.
XX Sequence 20 AA;
SQ Query Match 100.0%; Score 66; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19
RESULT 5
AAAY94321
ID AAAY94321 standard; peptide; 20 AA.
XX AC AAAY94321;
XX 11-AUG-2000 (first entry)
XX Mouse endostatin N-terminus.
XX Mouse; endothelial cell proliferation inhibitor; collagen XVIII;
XX angiogenesis inhibitor; anti-tumour; cytostatic; antipeptidic;
XX vasotropic; dermatological; ophthalmological; vulnary;
XX antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;
XX ocular angiogenic disease; atherosclerosis; scleroderma;
XX myocardial angiogenesis; telangiectasia; angiofibroma; wound granulation.
XX Mus musculus.
XX OS XX

PN WO200026368-A2.
XX 11-MAY-2000.
XX 01-NOV-1999; 99WO-US025605.
XX 30-OCT-1998; 98US-0106343P.
XX 20-MAY-1999; 99US-00315689.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX O'Reilly MS, Folkman MJ;
XX WPI; 2000-365617/31.
XX Novel endostatin capable of inhibiting endothelial cell proliferation and
XX angiogenesis, useful for treating angiogenesis-dependent cancers and as
XX birth control agents.
XX Disclosure; Page 2; 68pp; English.
XX The present sequence is the N-terminus of an endostatin, a potent and
XX specific protein inhibitor of endothelial cell proliferation in vitro and
XX angiogenesis in vivo. The sequence corresponds to a C-terminal fragment
XX of newly identified collagen type XVIII. Recombinant mouse endostatin (20
XX mg/kg) was administered subcutaneously to mice implanted with Lewis lung
XX carcinomas. There was tumour mass regression non-detectable levels after
XX 12 days of therapy due to the angiogenesis inhibitory activity of
XX endostatin. Thus the protein is useful for treatment of angiogenesis-
XX dependent cancers. The polynucleotide and polypeptide sequences of this
XX endostatin are useful for treating and diagnosis of tumours, ocular
XX angiogenic diseases, Osler-Webber syndrome, myocardial angiogenesis,
XX plaque neovascularisation, telangiectasia, haemophilic joints,
XX angiofibroma and wound granulation, for treatment of diseases related to
XX excessive or abnormal stimulation of endothelial cells e.g. intestinal
XX adhesions, atherosclerosis, scleroderma. The protein may also be useful
XX as a birth control agent by reducing or preventing uterine
XX vascularisation. The gene for endostatin may be isolated from cells or
XX tissue that express high levels of endostatin, eg. tumour cells, by
XX generating cDNA from mRNA using reverse transcriptase and then amplifying
XX the DNA sequence
XX Sequence 20 AA;
SQ Query Match 100.0%; Score 66; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19
RESULT 6
AAM48821
ID AAM48821 standard; peptide; 20 AA.
XX AC AAM48821;
XX 04-APR-2002 (first entry)
XX Endostatin N-terminal peptide.
XX Human; angiotatin; endostatin; angiogenesis; cancer; metastasis;
XX psoriasis; scleroderma; Crohn's disease; corneal disease; retinopathy;
XX arthritis; wound healing; Helicobacter pylori; peptic ulcer;
XX gene therapy; angiotatin antagonist; endostatin antagonist;
XX antiangiogenic; cytostatic; antiarthritic; antiinflammatory;
XX cerebroprotective; antidiabetic; virucide; antipyretic; vulnary;
XX gynaecological; cat scratch fever.
XX Unidentified.
XX OS XX

```

PN WO200193897-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 04-JUN-2001; 2001WO-US017947.  
 XX  
 PR 02-JUN-2000; 2000US-0203065P.  
 XX  
 PF 08-MAY-2001; 2001US-0289387P.  
 XX  
 PA (ENTR-) ENTREMED INC.  
 XX  
 PI Sim KL, Macdonald NU;  
 XX  
 XX WPI; 2002-130569/17.  
 DR  
 XX  
 XX  
 PT Regulating angiogenesis and treatment of angiogenesis-mediated diseases,  
 PT e.g. hemangioma, tumors or cancer, by administering a tropomyosin binding  
 PT compound or actin disrupting compound.  
 XX  
 XX Disclosure; Page 13; 95pp; English.  
 PS  
 XX The present invention relates to methods of regulating angiogenesis in an  
 CC individual by administering an angiogenesis regulating composition  
 CC comprising a tropomyosin binding compound or an actin disrupting  
 CC compound. The compositions are useful for treating diseases and processes  
 CC mediated by angiogenesis including haemangioma, solid tumours, blood  
 CC borne tumours, leukaemia, metastasis, Crohn's disease, coronary or  
 CC cerebral collaterals, arthritis, diabetic neovascularisation, macular  
 CC degeneration, wound healing, Helicobacter related diseases, ovulation,  
 CC menstruation, and cat scratch fever. The present sequence is a peptide  
 CC described in the exemplification of the invention  
 XX  
 XX Sequence 20 AA;  
 SQ  
 Query Match 100.0%; Score 66; DB 5; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPVLHLVALNTPL 13  
 DB 7 QPVLHLVALNTPL 19  
 RESULT 7  
 ADA43206  
 ID ADA43206 standard; peptide; 20 AA.  
 XX  
 AC ADA43206;  
 XX  
 XX 20-NOV-2003 (first entry)  
 DT  
 DE Mouse endostatin N-terminus.  
 XX  
 XX Mouse; endostatin; N-terminus; cytostatic; antipsoriatic; vulnery;  
 KW antiarteriosclerotic; angiogenesis-related disease;  
 KW angiogenesis-dependent cancer; birth control; psoriasis;  
 KW wound granulation; atherosclerosis; collagen XVIII.  
 XX  
 OS Mus sp.  
 XX  
 XX US2003087393-A1.  
 PN  
 PD 08-MAY-2003.  
 XX  
 XX 03-SEP-2002; 2002US-00232316.  
 PF  
 XX 16-OCT-1998; 98US-00174381.  
 PR  
 XX (OREI/) O'REILLY M S.  
 PA (FOLK/) FOLKMAN M J.  
 XX  
 XX O'reilly MS, Folkman MJ;  
 PI  
 XX

DR WPI; 2003-635177/60.  
 XX  
 PT New endostatin, useful for preparing a composition for treating an  
 PT angiogenesis-related disease, e.g. angiogenesis-dependent cancer,  
 PT psoriasis, wound granulation or atherosclerosis.  
 XX  
 XX Claim 3; Fig 5; 35pp; English.  
 PS  
 XX The invention relates to an isolated mouse endostatin (potent inhibitor  
 CC endothelial proliferation). Also include are a compound comprising an  
 CC isolated nucleic acid sequence encoding the endostatin protein (or an  
 CC isolated antibody capable of specifically binding to endostatin protein),  
 CC treating an angiogenesis-related disease, treating angiogenesis-dependent  
 CC cancer, a method of birth control, a composition comprising angiotatin  
 CC and endostatin and making the endostatin protein. The endostatin is  
 CC useful for preparing a composition for treating an angiogenesis-related  
 CC disease, e.g. angiogenesis-dependent cancer, psoriasis, wound granulation  
 CC or atherosclerosis. Microsequence analysis of the 20 kDa inhibitor of  
 CC capillary endothelial cell proliferation from the conditioned media  
 CC revealed identity to a C-terminal fragment of collagen XVIII. An 18-amino  
 CC acid N-terminal microsequence analysis of the purified inhibitor of 20  
 CC kDa inhibitor of capillary endothelial cell proliferation confirmed that  
 CC it was identical to a C-terminal fragment of the NCI domain. The  
 CC inhibitory fragment of collagen XVIII was named endostatin. The present  
 CC sequence is the N-terminus of the isolated endostatin.  
 XX  
 XX Sequence 20 AA;  
 SQ  
 Query Match 100.0%; Score 66; DB 7; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPVLHLVALNTPL 13  
 DB 7 QPVLHLVALNTPL 19  
 RESULT 8  
 ADM99393  
 ID ADM99393 standard; peptide; 20 AA.  
 XX  
 AC ADM99393;  
 XX  
 XX 29-JUL-2004 (first entry)  
 DT  
 DE Human endostatin N-terminal peptide.  
 XX  
 XX cytostatic; vulnery; vasotropic; antipsoriatic; dermatological;  
 KW cardiant; ophthalmological; antidiabetic; antiarthritic; antiulcer;  
 KW osteopathic; gynaecological; contraceptive; endostatin;  
 KW sucrose octasulfate; angiogenesis inhibitor;  
 KW endothelial proliferation inhibitor; endostatin biosynthesis;  
 KW anti-angiogenic compound; collagen type XVIII;  
 KW endothelial cell-related disorder; angiogenesis related disease; cancer;  
 KW haemangioma; leukaemia; metastasis; telangiectasia; psoriasis;  
 KW scleroderma; pyogenic granuloma; myocardial angiogenesis;  
 KW plaque neovascularisation; coronary collateral;  
 KW arteriovenous malformation; ischaemic limb angiogenesis; corneal disease;  
 KW rubeosis; neovascular glaucoma; diabetic retinopathy;  
 KW retrolental fibroplasia; arthritis; diabetic neovascularisation;  
 KW macular degeneration; wound healing; peptic ulcer; fracture; keloid;  
 KW vasculogenesis; haematopoiesis; ovulation; placentaion; menstruation;  
 KW birth control; human.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US2004091465-A1.  
 PN  
 PD 13-MAY-2004.  
 XX  
 XX 26-JUN-2003; 2003US-00607501.  
 PF  
 XX 26-JUN-2002; 2002US-0391630P.  
 PR



```

XX SQ Sequence 39 AA;
Query Match 100.0%; Score 66; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
DB 7 QPVLHLVALNTPL 19

RESULT 11
AAB71931
ID AAB71931 standard; protein; 42 AA.
XX AC AAB71931;
XX DT 10-MAY-2001 (first entry)
XX DE Murine sig-mEndo fusion protein N-terminal sequence.
XX KW Mouse; endostatin; antitumour; cytostatic; antiarthritic; antipsoriatic;
XX KW antidiabetic; ophthalmological; gene therapy; angiogenic inhibitor;
XX KW adenoviral vector; diabetic retinopathy; cardiovascular disease;
XX KW arthritis; psoriasis; cerebral oedema; intravascular coagulopathy;
XX KW lymphoma; leukaemia; sig-mEndo; fusion protein.
XX OS Mus sp.
XX PN WO200112830-A1.
XX PD 22-FEB-2001.
XX PF 11-AUG-2000; 2000WO-BP007865.
XX PR 13-AUG-1999; 99US-00373938.
XX PS (NOVS ) NOVARTIS AG.
XX PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX PI Hallenbeck PL, Chen CT;
XX DR WPI; 2001-202871/20.
XX PT Adenoviral vector for treating tumors and disorders associated with
XX PT angiogenesis, such as cancer, arthritis, and psoriasis, comprises a DNA
XX PT sequence encoding an angiogenic inhibitor, particularly endostatin.
XX PS Example 1; Fig 5; 59pp; English.
XX CC The present peptide was analysed in a study of the expression and
XX CC analysis of sig-mEndo, which comprises murine endostatin attached to the
XX CC murine Ig-kappa signal sequence. sig-mEndo was introduced into an
XX CC adenoviral vector. The adenoviral vector is useful for expressing
XX CC endostatin in a mammalian cell such as an A549 or Hep3B cell. It is
XX CC useful for treating other diseases and disorders associated with
XX CC angiogenesis, such as neovascular diseases of the eye, including diabetic
XX CC retinopathy, cardiovascular disease, arthritis, psoriasis, cerebral
XX CC oedema and intravascular coagulopathy (Kasabach-Merritt syndrome). The
XX CC vector inhibits, prevents or destroys the growth of tumours by preventing
XX CC the formation of blood vessels in tumours, such as lymphoma and leukaemia
XX SQ Sequence 42 AA;
Query Match 100.0%; Score 66; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
DB 30 QPVLHLVALNTPL 42

RESULT 12
AAM49504
ID AAM49504 standard; protein; 183 AA.
XX AC AAM49504;
XX DT 07-MAY-2002 (first entry)
XX DE Mouse endostatin protein.
XX KW Endostatin; murine; proliferation; blood vessel endothelium;
XX KW regeneration; tumour; blood vessel; treatment; amplification.
XX OS Mus sp.
XX FH Key
XX FT Misc-difference 181
XX FT /label= Ser, Phe
XX PN CN1177005-A.
XX PD 25-MAR-1998.
XX PF 10-SEP-1997; 97CN-00107112.
XX PR 10-SEP-1997; 97CN-00107112.
XX PA (XUGG/) XU G.
XX PI Xu G, Ren M, Xu L;
XX DR WPI; 2002-106746/15.
XX PT Gene clone of inhibitory factor for hyperplasia of inner blood vessel
XX PT cells in human body's real tumor, and its use in anti-tumor blood vessel
XX PT regeneration.
XX PS Disclosure; Page 4 (Disclosure); 6pp; Chinese.
XX CC This invention describes a novel preparation which inhibits the
XX CC proliferation of blood vessel endothelium and prevents the regeneration
XX CC activity of tumour blood vessels. The preparation can also be used as a
XX CC biological preparation in the treatment of tumours. This sequence
XX CC represents the murine endostatin protein described in the invention
XX SQ Sequence 183 AA;
Query Match 100.0%; Score 66; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
DB 7 QPVLHLVALNTPL 19

RESULT 13
AAV18409
ID AAV18409 standard; protein; 184 AA.
XX AC AAV18409;
XX DT 24-AUG-1999 (first entry)
XX DE Endostatin protein sequence.
XX KW BM1; anti-angiogenic peptide; endostatin; angiogenesis-dependent cancer;
XX KW benign tumour; rheumatoid arthritis; psoriasis; ocular angiogenesis;
XX KW Osler-Webber Syndrome; myocardial angiogenesis; angiofibroma; cancer;
XX KW plaque neovascularisation; telangiectasis; atherosclerosis; scleroderma;
XX KW dialysis graft vascular access stenosis; renal cancer; therapy.
XX
```

OS Mus sp.  
 XX WO929855-A1.  
 PN  
 XX  
 PD 17-JUN-1999.  
 XX  
 XX  
 PF 08-DEC-1998; 98WO-US026057.  
 XX  
 XX 08-DEC-1997; 97US-0067888P.  
 PR 22-APR-1998; 98US-0082663P.  
 PR 16-NOV-1998; 98US-0108536P.  
 XX  
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 XX  
 XX Sukhatme VP;  
 PI  
 XX  
 XX WPI: 1999-385604/32.  
 DR N-PSDB; AAX79949.  
 DR  
 XX  
 PT Mutant endostatin having anti-angiogenic activity.  
 XX  
 XX Claim 31; Fig 2; 105pp; English.  
 PS  
 XX This sequence is the mouse endostatin. The invention relates to a the  
 CC mutant endostatin (EM), which has anti-angiogenic activity, and is  
 CC designated EMI. Compositions comprising EMI or fusion proteins comprising  
 CC EMI, are useful for treating diseases characterised by angiogenic  
 CC activity, such as angiogenesis-dependent cancers, benign tumours,  
 CC rheumatoid arthritis, psoriasis, ocular angiogenesis, Osler-Webber  
 CC Syndrome, myocardial angiogenesis, plaque neovascularisation,  
 CC telangiectasia, haemophilic joints, angiofibroma, wound granulation,  
 CC intestinal adhesions, atherosclerosis, scleroderma, hypertrophic scars,  
 CC cat scratch disease, Helicobacter pylori ulcers, dialysis graft vascular  
 CC access stenosis, contraception and obesity. In particular, the diseases  
 CC treatable by EMI comprise cancer, especially renal cancer. The methods  
 CC provide a means for introducing EMI into mammalian cells via gene  
 CC therapy, for production of EMI via recombinant means, as well as  
 CC recombinant production of the EMI protein. EMI performs as well or better  
 CC than whole endostatin. Use of EMI is advantageous for treatment of  
 CC angiogenic diseases in that increasingly smaller peptides are more potent  
 CC on a weight basis, and may be able to better penetrate tissues  
 XX  
 XX SQ Sequence 184 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.00071;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPVLHLVALNTPL 13  
 Db | | | | | | | | | |  
 7 QPVLHLVALNTPL 19  
 RESULT 14  
 AAY08689  
 ID AAY08689 standard; protein; 184 AA.  
 XX  
 XX AC AAY08689;  
 XX  
 XX DT 10-AUG-1999 (first entry)  
 XX  
 XX DE Murine endostatin protein fragment.  
 DE  
 XX Plasmidogen; murine; angiotensin; endostatin; gene therapy; vector;  
 KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;  
 KW tumour growth; solid tumour; diabetic retinopathy; retina.  
 XX  
 XX OS Mus sp.  
 XX WO9926480-A1.  
 PN  
 XX  
 XX 03-JUN-1999.  
 PD

PF 20-NOV-1998; 98WO-US024950.  
 XX  
 XX 20-NOV-1997; 97US-00975424.  
 XX  
 XX (GENE-) GENETIX PHARM INC.  
 PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
 XX  
 XX Leboulch P, Pawliuk RJ, Bachelot T;  
 PI  
 XX WPI: 1999-357696/30.  
 DR N-PSDB; AAX77715.  
 DR  
 XX  
 XX PT Anti-angiogenic gene therapy vectors.  
 XX  
 XX PS Disclosure; Fig 6; 83pp; English.  
 XX  
 XX This invention describes a novel viral gene therapy vector comprising a  
 CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from  
 CC human or murine angiotensin, human or murine endostatin and angiogenesis-  
 CC inhibiting fusions and fragments, where the viral vector is sufficiently  
 CC attenuated for use in human gene therapy. The products of the invention  
 CC have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological  
 CC activity. The vector is used in gene therapy for inhibiting tumour growth  
 CC in humans harbouring a solid tumour. The vector expresses an anti-  
 CC angiogenic polypeptide. An additional use comprises treatment of diabetic  
 CC retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis  
 CC in the vicinity of the retina. The vector is administered to cells ex  
 CC vivo and then administered to the patient  
 XX  
 XX SQ Sequence 184 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.00071;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QPVLHLVALNTPL 13  
 Db | | | | | | | | | |  
 7 QPVLHLVALNTPL 19  
 RESULT 15  
 AAY70258  
 ID AAY70258 standard; protein; 184 AA.  
 XX  
 XX AC AAY70258;  
 XX  
 XX DT 06-JUN-2000 (first entry)  
 XX  
 XX DE Murine angiogenesis inhibitor, endostatin.  
 XX  
 XX KW Murine; immunoglobulin Fc fragment; endostatin; immunofusin;  
 KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;  
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;  
 KW vasotropic; vulnary; treatment; antiarteriosclerosis; tumour;  
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW wound granulation; keloid scar; gene therapy.  
 XX  
 XX OS Mus musculus.  
 XX  
 XX PN WO200011033-A2.  
 XX  
 XX PD 02-MAR-2000.  
 XX  
 XX PF 25-AUG-1999; 99WO-US019329.  
 XX  
 XX PR 25-AUG-1998; 98US-0097883P.  
 XX  
 XX PA (LEXI-) LEXINGEN PHARM CORP.  
 XX  
 XX PI Lo K, Li Y, Gillies SD;  
 XX

DR WPI: 2000-237616/20.  
DR N-PSDB; AAZ51299.  
XX  
PT Novel fusion protein of angiostatin or endostatin and an immunoglobulin  
PT FC region, useful for treating conditions mediated by angiogenesis, such  
PT as rheumatoid arthritis, tumors and macular degeneration.  
XX  
PS Example 5; Page 48-49; 68pp; English.  
XX  
CC The patent discloses a DNA molecule encoding a fusion protein comprising  
CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis  
CC inhibitor selected from angiostatin, endostatin, a plasminogen fragment  
CC having angiostatin activity, a collagen XVIII fragment having endostatin  
CC activity, or combinations of them. The fusion protein (immunofusin) is  
CC used to inhibit angiogenesis and to treat diseases or conditions mediated  
CC by angiogenesis. Conditions that may be treated include solid tumours,  
CC blood born tumours, tumour metastasis, benign tumours including  
CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic  
CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases  
CC e.g. diabetic retinopathy, retinopathy of prematurity, macular  
CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental  
CC fibroplasia, rubrosis and Osler-Webber syndrome; myocardial angiogenesis,  
CC plaque neovascularisation, telangiectasia, haemophilic joints,  
CC angiofibroma, wound granulation, and excessive or abnormal stimulation of  
CC endothelial cells, intestinal cells, atherosclerosis, sclerodermal and  
CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in  
CC gene therapy. The present sequence is a murine endostatin used in the  
CC construction of immunofusin containing murine immunoglobulin Fc fragment  
XX  
SQ Sequence 184 AA;  
  
Query Match 100.0%; Score 66; DB 3; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.00071;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QPVLHLVALNTEPL 13  
| | | | | | | | | | | | | | |  
Db 7 QPVLHLVALNTEPL 19  
  
Search completed: March 10, 2005, 11:06:51  
Job time : 67.0263 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:04:09 ; Search time 26.3421 Seconds  
(without alignments)  
36.840 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66

Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	3	US-09-385-442-30
2	66	100.0	20	2	US-08-740-168A-1
3	66	100.0	20	3	US-09-349-429-1
4	66	100.0	20	3	US-09-315-689-1
5	66	100.0	20	4	US-09-174-282-1
6	66	100.0	20	4	US-09-154-302-1
7	66	100.0	20	4	US-09-174-381-1
8	66	100.0	20	4	US-09-405-499-1
9	66	100.0	184	4	US-09-938-391A-13
10	66	100.0	185	3	US-08-985-526-36
11	66	100.0	191	3	US-09-561-500-13
12	66	100.0	191	3	US-09-561-108-13
13	66	100.0	191	4	US-09-561-526-13
14	66	100.0	191	4	US-09-561-499-13
15	66	100.0	191	4	US-09-998-831-13
16	66	100.0	191	4	US-09-561-005-13
17	66	100.0	191	1	US-08-159-784-2
18	62	93.9	178	3	US-09-315-689-5
19	62	93.9	182	3	US-09-561-500-14
20	62	93.9	182	3	US-09-561-108-14
21	62	93.9	182	3	US-09-315-689-3
22	62	93.9	182	4	US-09-561-526-14
23	62	93.9	182	4	US-09-561-499-14
24	62	93.9	182	4	US-09-998-831-14
25	62	93.9	182	4	US-09-561-005-14
26	62	93.9	183	3	US-09-206-059-2
27	62	93.9	183	4	US-09-938-391A-12

28	62	93.9	184	4	US-09-938-391A-11	Sequence 11, Appl
29	62	93.9	684	4	US-09-961-403-5	Sequence 5, Appli
30	62	93.9	1516	4	US-09-949-016-8209	Sequence 8209, Ap
31	48	72.7	184	4	US-09-938-391A-4	Sequence 4, Appli
32	48	72.7	230	4	US-09-938-391A-2	Sequence 2, Appli
33	44	66.7	15	4	US-09-231-077D-12	Sequence 12, Appl
34	44	66.7	15	4	US-09-231-077D-13	Sequence 13, Appl
35	44	66.7	181	4	US-09-589-483-20	Sequence 20, Appl
36	44	66.7	191	1	US-08-159-784-3	Sequence 3, Appli
37	43	65.2	1867	4	US-09-824-574-5	Sequence 5, Appli
38	41	62.1	955	4	US-09-252-991A-18882	Sequence 18882, A
39	40	60.6	413	4	US-09-724-623-87	Sequence 87, Appl
40	38	57.6	2311	3	US-08-934-386-9	Sequence 9, Appli
41	37	56.1	143	4	US-09-270-767-40331	Sequence 40331, A
42	37	56.1	143	4	US-09-270-767-55547	Sequence 55547, A
43	37	56.1	187	1	US-08-471-058-17	Sequence 17, Appl
44	37	56.1	187	3	US-08-471-057-17	Sequence 17, Appl
45	37	56.1	187	4	US-08-470-865-17	Sequence 17, Appl

## ALIGNMENTS

### RESULT 1

US-09-385-442-30  
; Sequence 30, Application US/09385442  
; Patent No. 6200954

GENERAL INFORMATION:

APPLICANT: Ge, Ruowen

APPLICANT: Kini, R. Manjunatha

TITLE OF INVENTION: Small Peptides Having Potent Anti-Angiogenic Activity

FILE REFERENCE: 1781-170P

CURRENT APPLICATION NUMBER: US/09/385,442

CURRENT FILING DATE: 1999-08-30

EARLIER APPLICATION NUMBER: 60/099,313

EARLIER FILING DATE: 1999-09-04

NUMBER OF SEQ ID NOS: 50

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 30

LENGTH: 13

TYPE: PRT

ORGANISM: mammalian

FEATURE:

OTHER INFORMATION: Endo-2

US-09-385-442-30

Query Match 100.0%; Score 66; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13

Db 1 QPVLHLVALNTPL 13

### RESULT 2

US-08-740-168A-1

; Sequence 1, Application US/08740168A

Patent No. 5854205

GENERAL INFORMATION:

APPLICANT: O'Reilly, Michael

APPLICANT: Folkman, M. Judah

TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions

TITLE OF INVENTION: and Methods

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: Jones & Askew, LLP

STREET: 191 Peachtree, 37th Floor

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,168A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
; US-08-740-168A-1

Query Match 100.0%; Score 66; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 3
US-09-349-429-1
; Sequence 1, Application US/09349429
; Patent No. 6174861
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/349,429
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,168
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
; US-09-349-429-1

Query Match 100.0%; Score 66; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 4
US-09-315-689-1
; Sequence 1, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: murine
; US-09-315-689-1

Query Match 100.0%; Score 66; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 5
US-09-174-282-1
; Sequence 1, Application US/09174282
; Patent No. 6544758
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/174,282  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/740,168  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714  
REFERENCE/DOCKET NUMBER: 05213-0223  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-818-3700  
TELEFAX: 404-818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Murine  
TISSUE TYPE: Collagen  
US-09-174-282-1

Query Match 100.0%; Score 66; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 7 QPVLHLVALNTPL 19

RESULT 6  
US-09-154-302-1  
Sequence 1, Application US/09154302  
Patent No. 6630448  
GENERAL INFORMATION:  
APPLICANT: O'Reilly, Michael  
APPLICANT: Folkman, M. Judah  
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions  
TITLE OF INVENTION: and Methods  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones & Askew, LLP  
STREET: 191 Peachtree, 37th Floor  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,302  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/740,168  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714

REFERENCE/DOCKET NUMBER: 05213-0223  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-818-3700  
TELEFAX: 404-818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Murine  
TISSUE TYPE: Collagen  
US-09-154-302-1

Query Match 100.0%; Score 66; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 7 QPVLHLVALNTPL 19

RESULT 7  
US-09-174-381-1  
Sequence 1, Application US/09174381  
Patent No. 674685  
GENERAL INFORMATION:  
APPLICANT: O'Reilly, Michael  
APPLICANT: Folkman, M. Judah  
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions  
TITLE OF INVENTION: and Methods  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones & Askew, LLP  
STREET: 191 Peachtree, 37th Floor  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/174,381  
FILING DATE: 16-Oct-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/740,168  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714  
REFERENCE/DOCKET NUMBER: 05213-0223  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-818-3700  
TELEFAX: 404-818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

```
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-174-381-1
    Query Match      100.0%; Score 66; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 9.3e-05;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 8
US-09-405-499-1
; Sequence 1, Application US/09405499
; Patent No. 6764995
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Endostatin Protein and Fragments Thereof
; FILE REFERENCE: 05213-0640
; CURRENT APPLICATION NUMBER: US/09/405,499
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-405-499-1
    Query Match      100.0%; Score 66; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 9.3e-05;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 9
US-09-938-391A-13
; Sequence 13, Application US/09938391A
; Patent No. 6803211
; GENERAL INFORMATION:
; APPLICANT: Pfizer, Inc. and Pfizer Products, Inc.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING DISORDERS
; TITLE OF INVENTION: INVOLVING ANGIOGENESIS
; FILE REFERENCE: 3153.00234/PCI0790A
; CURRENT APPLICATION NUMBER: US/09/938,391A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US PROV NO. 60/227,924
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 184
; TYPE: PRT
; ORGANISM: MURINE
US-09-938-391A-13
    Query Match      100.0%; Score 66; DB 4; Length 184;
    Best Local Similarity 100.0%; Pred. No. 0.001;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 10
US-08-985-526-36
; Sequence 36, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mikson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorrow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-985-526-36
    Query Match      100.0%; Score 66; DB 3; Length 185;
    Best Local Similarity 100.0%; Pred. No. 0.001;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 8 QPVLHLVALNTPL 20

RESULT 11
US-09-561-500-13
; Sequence 13, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-500-13
```

```
Query Match      100.0%; Score 66; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 12
US-09-561-108-13
; Sequence 13, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-108-13

Query Match      100.0%; Score 66; DB 3; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 13
US-09-561-526-13
; Sequence 13, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-526-13

Query Match      100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26
```

```
RESULT 14
US-09-561-499-13
; Sequence 13, Application US/09561499
; Patent No. 6524583
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/09/561,499
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-499-13

Query Match      100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 15
US-09-998-831-13
; Sequence 13, Application US/09998831
; Patent No. 6676941
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-998-831-13

Query Match      100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

Search completed: March 10, 2005, 11:11:30
Job time : 27.3421 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 10:55:58 ; Search time 55.8684 Seconds  
(without alignments)  
76.150 Million cell updates/sec

Title: US-09-766-412-29  
Perfect score: 59  
Sequence: 1 SPHNSYIVLPI 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	11	4	Aab74255 Exemplary
2	59	100.0	11	4	Aab80861 Angicostat
3	59	100.0	11	5	Abg97538 Antiangio
4	40	67.8	638	6	Abu20188 Protein e
5	39	66.1	48	4	Aab49808 Murine en
6	39	66.1	50	4	Aab35587 Antiangio
7	39	66.1	58	5	Aam49504 Mouse end
8	39	66.1	183	5	AAM49504 Mouse end
9	39	66.1	184	2	AAY18409 Endostati
10	39	66.1	184	2	AAY08689 Murine en
11	39	66.1	184	3	AAY70258 Murine an
12	39	66.1	184	4	AAB49380 Murine en
13	39	66.1	184	5	ABG31793 Human end
14	39	66.1	184	7	ABU64290 Murine en
15	39	66.1	191	3	AAB28398 Murine en
16	39	66.1	191	5	AAY77950 Amino aci
17	39	66.1	201	6	ABJ25793 Aspergill
18	39	66.1	201	6	ABJ26393 Aspergill
19	39	66.1	207	4	Aae02031 Murine en
20	39	66.1	207	4	AAB71930 Murine en
21	39	66.1	207	5	ABb79902 Mouse end
22	39	66.1	207	8	Adn00602 Murine en
23	39	66.1	207	8	Ado43904 Amino aci
24	39	66.1	218	2	Aay08691 Murine ge
25	39	66.1	231	7	Adh87724 Enterococ

26	39	66.1	455	4	AAU33460	Aau33460 Enterococ
27	39	66.1	466	6	ABU44591	Abu44591 Protein e
28	39	66.1	472	4	AAU35185	Aau35185 Enterococ
29	39	66.1	472	6	ABU14632	Abu14632 Protein e
30	39	66.1	580	2	AAy08692	Aay08692 Murine ge
31	39	66.1	684	2	AAy25114	Aay25114 Mouse alp
32	39	66.1	1288	2	AAW26328	Aaw26328 Mouse alp
33	39	66.1	1288	2	AAW92297	Aaw92297 Mouse alp
34	39	66.1	1315	8	ABO84585	AbO84585 Mouse can
35	39	66.1	1527	8	ABO84584	AbO84584 Mouse can
36	39	66.1	1987	4	ABB61083	ABb61083 Drosophil
37	38.5	65.3	818	8	ADS43599	AdS43599 Bacterial
38	38	64.4	441	7	ABO83463	AbO83463 Pseudomon
39	38	64.4	628	6	ABU21359	Abu21359 Protein e
40	37	62.7	123	6	ADB09837	AdB09837 Alloiococ
41	37	62.7	227	6	ABU33803	Abu33803 Protein e
42	37	62.7	245	6	ABU34059	Abu34059 Protein e
43	37	62.7	264	6	ABU36425	Abu36425 Protein e
44	37	62.7	265	6	ABU34704	Abu34704 Protein e
45	37	62.7	289	5	AAO17569	Aao17569 M catarrh

## ALIGNMENTS

RESULT 1  
AAB74255  
ID AAB74255 standard; peptide; 11 AA.  
XX  
AC AAB74255;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Exemplary anti-angiogenic peptide #29.  
XX  
KW Anti-angiogenesis; plasminogen; VEGF; vascular endothelial growth factor;  
KW FLT-1; tumour; metastasis; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200118030-A2.  
XX  
PD 15-MAR-2001.  
XX  
PF 01-SEP-2000; 2000WO-SG000131.  
XX  
PR 03-SEP-1999; 99SG-00004310.  
XX  
(UYSI-) UNIV SINGAPORE NAT.  
Ge R, Kini RM;  
WPI; 2001-257785/26.  
Peptides comprising a portion of a protein selected from plasminogen, endostatin, VEGF, FLT-1 and KDR/FLK-1 are useful for treating primary tumor growth.  
Claim 7; Page 18; 34pp; English.  
The present invention relates to anti-angiogenesis peptides from a portion of a selected from plasminogen, endostatin, VEGF or FLT-1. The invention is used to prevent or treat primary tumour growth or metastasis or undesired angiogenesis  
Sequence 11 AA;  
Query Match 100.0%; Score 59; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SPHNSYIVLPI 11  
|||||

```

Db      1 SPHNSYIVLPI 11

RESULT 2
AAB80861
ID AAB80861 standard; peptide; 11 AA.
XX
AC AAB80861;
XX
XX
DT 29-MAY-2001 (first entry)
XX
DE Angiostatin-derived peptide #18.
XX
KW Cytostatic; angiogenesis inhibitor; tumour metastasis; tumour growth.
XX
OS Unidentified.
XX
XX
PN US6200954-B1.
XX
XX
PD 13-MAR-2001.
XX
XX
PF 30-AUG-1999; 99US-00385442.
XX
PR 04-SEP-1998; 98US-0099313P.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Ge R, Kini RM;
XX
DR WPI; 2001-234520/24.
XX
XX
PT Peptides having potent anti-angiogenic activity, useful for the treatment
PT of tumors, comprise a portion of a FLT-1 protein.
XX
PS Disclosure; Col 15-16; 21pp; English.
XX
XX
CC The present sequence is a peptide which is effective in inhibiting
CC undesirable angiogenesis. Angiogenesis is the process of new blood vessel
CC formation from pre-existing vessels. Inappropriate angiogenesis is
CC associated with various pathological conditions including solid tumour
CC growth and metastasis. The present peptide can be used to prevent
CC tumour metastasis or inhibit the growth of a primary tumour
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 59; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11
   |||||
Db 1 SPHNSYIVLPI 11

RESULT 3
ABG97538
ID ABG97538 standard; peptide; 11 AA.
XX
AC ABG97538;
XX
XX
DT 16-DEC-2002 (first entry)
XX
DE Antiangiogenic peptide, Endo-1.
XX
KW Angiogenesis; plasminogen; endostatin; Endo-; VEGF;
KW vascular endothelial growth factor; FLT-1; kinase insert domain; FLK-1;
KW angiogenic inhibitor; Angio-; endothelial cell; proliferation;
KW tumour growth; blood vessel formation; migration; tubule structure;
KW embryonic development; wound healing; tumour metastasis;
KW rheumatoid arthritis; psoriasis; anticancer; therapy;
KW antiangiogenic therapy; mitogen; tyrosine kinase receptor.
XX
OS Mammalia.

Db      1 SPHNSYIVLPI 11

RESULT 4
ABU20188
ID ABU20188 standard; protein; 638 AA.
XX
AC ABU20188;
XX
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #5715.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Borrelia cepacia.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.

```

```

XX US2002103129-A1.
PN
XX
XX 01-AUG-2002.
XX
XX 22-JAN-2001; 2001US-00766412.
XX
XX 04-SEP-1998; 98US-0099313P.
PR 30-AUG-1999; 99US-00385442.
XX
XX (GERR/) GE R.
PA (KINI/) KINI R M.
XX
XX Ge R, Kini RM;
XX
XX WPI; 2002-749595/81.
XX
XX Novel peptide for preventing/treating undesired angiogenesis, has a
PT portion of a protein e.g. plasminogen, endostatin, and potent
PT antiangiogenic activity and endothelial cell proliferation inhibition
PT activity.
XX
XX Claim 9; Page 9; 24pp; English.
XX
XX The invention discloses a peptide comprising a portion of a protein e.g.
CC plasminogen, endostatin (Endo-1-4), vascular endothelial growth factor
CC (VEGF), FLT-1 or kinase insert domain containing receptor/FLK-1. These
CC peptides are able to act as potent angiogenic inhibitors (Angio-1-5)
CC inhibiting endothelial cell proliferation and retarding tumour growth.
CC Angiogenesis is the process of new blood vessel formation from pre-
CC existing vessel, involving endothelial cell proliferation, migration and
CC assembly into tubule structures. It plays an important role in normal
CC physiological functions such as embryonic development and wound healing.
CC Inappropriate angiogenesis is also associated with various pathological
CC conditions including tumour growth and metastasis, rheumatoid arthritis
CC and psoriasis. Anticancer therapy by inhibiting tumour angiogenesis is
CC called antiangiogenic therapy. VEGF is an endothelial specific mitogen
CC that functions through two high affinity tyrosine kinase receptors, FLT-1
CC and FLK-1 (not defined). Protein-protein interactions are crucial to many
CC physiological and pharmacological processes and the peptides disclosed
CC interfere with these interactions. The peptides are useful for preventing
CC or treating undesired angiogenesis and primary tumour growth or
CC metastasis. The sequences presented in ABG97510-ABG97559 are examples of
CC antiangiogenic peptides
XX
XX Sequence 11 AA;

Query Match 100.0%; Score 59; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11
   |||||
Db 1 SPHNSYIVLPI 11

RESULT 4
ABU20188
ID ABU20188 standard; protein; 638 AA.
XX
AC ABU20188;
XX
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #5715.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Borrelia cepacia.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.

```



XX 21-MAR-2002; 2002WO-US0093107.  
XX  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342323P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA24058.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 48112; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 638 AA;  
SQ  
Query Match 67.8%; Score 40; DB 6; Length 638;  
Best Local Similarity 63.6%; Pred. No. 1.1e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SPHNSYIVLPI 11  
|||:|||||  
DB 151 TPLSHLVLP 161  
RESULT 5  
AAB49808  
ID AAB49808 standard; protein; 48 AA.  
XX  
XX AAB49808;  
XX  
XX 02-MAR-2001 (first entry)  
DT

XX Murine endostatin peptide fragment SEQ ID NO: 21.  
XX  
XX Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken; cancer;  
XX inflammation; angiogenesis-dependent disease.  
XX  
XX Mus musculus.  
XX  
XX WO200067771-A1.  
XX  
XX 16-NOV-2000.  
XX  
XX 02-MAY-2000; 2000WO-US012063.  
XX  
XX 06-MAY-1999; 99US-0132907P.  
XX 14-JUL-1999; 99US-00353333.  
XX  
XX (BURN-) BURNHAM INST.  
XX  
XX Vuori K;  
XX  
XX WPI; 2001-040937/05.  
XX  
XX Endostatin peptide comprising at least four endostatin amino acid  
PT residues are e.g. angiogenesis inhibitors for treating cancer and  
PT diabetic retinopathy.  
XX  
XX Claim 10; Page 124-125; 146pp; English.  
XX  
XX The present invention provides endostatin peptides which can be used in  
CC the modulation of angiogenesis. This is useful in the treatment of  
CC cancers, inflammation, rheumatoid arthritis, chronic articular  
CC rheumatism, psoriasis, disorders associated with inopportune invasion of  
CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
CC of prematurity, macular degeneration, corneal graft rejection,  
CC retrolental fibroplasia, rubeosis, capillary proliferation in  
CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
CC neovascularisation, telangiectasia, haemophilic joints and wound  
CC granulation. In addition, the peptides can be used as birth control  
CC agents  
XX  
XX Sequence 48 AA;  
SQ  
Query Match 66.1%; Score 39; DB 4; Length 48;  
Best Local Similarity 81.8%; Pred. No. 9.2;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SPHNSYIVLPI 11  
|||:|||||  
DB 33 SCHNSYIVLCI 43  
RESULT 6  
AAB35587  
ID AAB35587 standard; peptide; 50 AA.  
XX  
XX AAB35587;  
XX  
XX 14-FEB-2001 (first entry)  
XX  
XX Antiangiogenic pentapeptide IV.  
XX  
XX Antiangiogenic; angiogenesis; cancer; endostatin.  
XX  
XX Synthetic.  
XX  
XX WO200063249-A1.  
XX  
XX 26-OCT-2000.  
XX  
XX 11-APR-2000; 2000WO-EP003236.  
XX



PS Disclosure; Page 4 (Disclosure); 6pp; Chinese.

XX This invention describes a novel preparation which inhibits the proliferation of blood vessel endothelium and prevents the regeneration of tumour blood vessels. The preparation can also be used as a biological preparation in the treatment of tumours. This sequence represents the murine endostatin protein described in the invention

XX Sequence 183 AA;

SQ Query Match 66.1%; Score 39; DB 5; Length 183;  
Best Local Similarity 81.8%; Pred. No. 41;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | |  
Db 164 SCHNSYIVLCI 174

RESULT 9  
AAY18409  
ID AAY18409 standard; protein; 184 AA.  
XX  
AC AAY18409;  
XX  
DT 24-AUG-1999 (first entry)  
XX  
DE Endostatin protein sequence.  
XX  
KW EM1; anti-angiogenic peptide; endostatin; angiogenesis-dependent cancer; benign tumour; rheumatoid arthritis; psoriasis; ocular angiogenesis; Osler-Webber Syndrome; myocardial angiogenesis; angiofibroma; cancer; plaque neovascularisation; telangiectasia; atherosclerosis; scleroderma; dialysis graft vascular access stenosis; renal cancer; therapy.

XX  
OS Mus sp.  
XX  
PN WO929855-A1.  
XX  
PD 17-JUN-1999.  
XX  
PF 08-DEC-1998; 98WO-US026057.  
XX  
PR 08-DEC-1997; 97US-0067888P.  
PR 22-APR-1998; 98US-0082663P.  
PR 16-NOV-1998; 98US-0108536P.  
XX  
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
XX  
PI Sukhatme VP;  
XX  
DR WPI; 1999-385604/32.  
DR N-PSDB; AAX79949.  
XX  
PT Mutant endostatin having anti-angiogenic activity.  
XX  
PS Claim 31; Fig 2; 105pp; English.  
XX  
CC This sequence is the mouse endostatin. The invention relates to a the mutant endostatin (EM1), which has anti-angiogenic activity, and is designated EM1. Compositions comprising EM1 or fusion proteins comprising EM1, are useful for treating diseases characterised by angiogenic activity, such as angiogenesis-dependent cancers, benign tumours, rheumatoid arthritis, psoriasis, ocular angiogenesis, Osler-Webber Syndrome, myocardial angiogenesis, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, wound granulation, intestinal adhesions, atherosclerosis, scleroderma, hypertrophic scars, cat scratch disease, Helicobacter pylori ulcers, dialysis graft vascular access stenosis, contraception and obesity. In particular, the diseases treatable by EM1 comprise cancer, especially renal cancer. The methods provide a means for introducing EM1 into mammalian cells via gene therapy, for production of EM1 via recombinant means, as well as recombinant production of the EM1 protein. EM1 performs as well or better

CC than whole endostatin. Use of EM1 is advantageous for treatment of angiogenic diseases in that increasingly smaller peptides are more potent on a weight basis, and may be able to better penetrate tissues

XX Sequence 184 AA;

SQ Query Match 66.1%; Score 39; DB 2; Length 184;  
Best Local Similarity 81.8%; Pred. No. 42;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | |  
Db 164 SCHNSYIVLCI 174

RESULT 10  
AAY08689  
ID AAY08689 standard; protein; 184 AA.  
XX  
AC AAY08689;  
XX  
DT 10-AUG-1999 (first entry)  
XX  
DE Murine endostatin protein fragment.  
XX  
KW Plasminogen; murine; angioostatin; endostatin; gene therapy; vector; anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology; tumour growth; solid tumour; diabetic retinopathy; retina.

XX  
OS Mus sp.  
XX  
PN WO926480-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 20-NOV-1998; 98WO-US024950.  
XX  
PR 20-NOV-1997; 97US-00975424.  
XX  
PA (GENE-) GENETIX PHARM INC.  
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.  
XX  
PI Leboulch P, Pawliuk RJ, Bachelot T;  
XX WPI; 1999-357696/30.  
DR N-PSDB; AAX77715.  
XX  
PT Anti-angiogenic gene therapy vectors.  
XX  
PS Disclosure; Fig 6; 83pp; English.  
XX  
CC This invention describes a novel viral gene therapy vector comprising a nucleic acid molecule encoding an anti-angiogenic polypeptide chosen from human or murine angioostatin, human or murine endostatin and angiogenesis-inhibiting fusions and fragments, where the viral vector is sufficiently attenuated for use in human gene therapy. The products of the invention have anti-angiogenic, cytostatic, anti-diabetic and ophthalmological activity. The vector is used in gene therapy for inhibiting tumour growth in humans harbouring a solid tumour. The vector expresses an anti-angiogenic polypeptide. An additional use comprises treatment of diabetic retinopathy, where the anti-angiogenic polypeptide inhibits angiogenesis in the vicinity of the retina. The vector is administered to cells ex vivo and then administered to the patient

XX Sequence 184 AA;

SQ Query Match 66.1%; Score 39; DB 2; Length 184;  
Best Local Similarity 81.8%; Pred. No. 42;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | |  
Db 164 SCHNSYIVLCI 174

RESULT 11  
 AAY70258  
 ID AAY70258 standard; protein; 184 AA.  
 XX  
 AC AAY70258;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Murine angiogenesis inhibitor, endostatin.  
 XX  
 KW Murine; immunoglobulin Fc fragment; endostatin; immunofusin;  
 KW angiogenesis; inhibitor; cytosolic; antirheumatoid; antiarthritic;  
 KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;  
 KW vasotropic; vulnary; treatment; antiarteriosclerosis; tumour;  
 KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW wound granulation; keloid scar; gene therapy.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200011033-A2.  
 XX  
 PD 02-MAR-2000.  
 XX  
 PF 25-AUG-1999; 99WO-US019329.  
 XX  
 PR 25-AUG-1998; 98US-0097883P.  
 XX  
 PA (LEXI-) LEXINGEN PHARM CORP.  
 XX  
 PI Lo K, Li Y, Gillies SD;  
 XX  
 DR WPI; 2000-237616/20.  
 DR N-PSDB; AA251299.  
 XX  
 PT Novel fusion protein of angiotatin or endostatin and an immunoglobulin  
 FC region, useful for treating conditions mediated by angiogenesis, such  
 PT as rheumatoid arthritis, tumors and macular degeneration.  
 XX  
 PS Example 5; Page 48-49; 68pp; English.  
 XX  
 CC The patent discloses a DNA molecule encoding a fusion protein comprising  
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis  
 CC inhibitor selected from angiotatin, endostatin, a plasminogen fragment  
 CC having angiotatin activity, a collagen XVII fragment having endostatin  
 CC activity, or combinations of them. The fusion protein (immunofusin) is  
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated  
 CC by angiogenesis. Conditions that may be treated include solid tumours,  
 CC blood born tumours, tumour metastasis, benign tumours including  
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic  
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases  
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular  
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental  
 CC fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,  
 CC plaque neovascularisation, telangiectasia, haemophilic joints.  
 CC angiobroma, wound granulation, and excessive or abnormal stimulation of  
 CC endothelial cells, intestinal cells, atherosclerosis, sclerodermal and  
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in  
 CC gene therapy. The present sequence is a murine endostatin used in the  
 CC construction of immunofusin containing murine immunoglobulin Fc fragment  
 XX  
 SQ Sequence 184 AA;  
 Query Match 66.1%; Score 39; DB 3; Length 184;  
 Best Local Similarity 81.8%; Pred. No. 42;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLPI 11  
 | | | | | | | |  
 Db 164 SCHNSYIVLCI 174

RESULT 12  
 AAB49380  
 ID AAB49380 standard; protein; 184 AA.  
 XX  
 AC AAB49380;  
 XX  
 DT 02-MAR-2001 (first entry)  
 XX  
 DE Murine endostatin SEQ ID NO: 4.  
 XX  
 KW Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken; cancer;  
 KW inflammation; angiogenesis-dependent disease.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200067771-A1.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 02-MAY-2000; 2000WO-US012063.  
 XX  
 PR 06-MAY-1999; 99US-0132907P.  
 PR 14-JUL-1999; 99US-00353333.  
 XX  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Vuori K;  
 XX  
 DR WPI; 2001-040937/05.  
 DR N-PSDB; AAC88290.  
 XX  
 PT Endostatin peptide comprising at least four endostatin amino acid  
 PT residues are e.g. angiogenesis inhibitors for treating cancer and  
 PT diabetic retinopathy.  
 XX  
 PS Disclosure; Fig 1; 146pp; English.  
 XX  
 CC The present invention provides endostatin peptides which can be used in  
 CC the modulation of angiogenesis. This is useful in the treatment of  
 CC cancers, inflammation, rheumatoid arthritis, chronic articular  
 CC rheumatism, psoriasis, disorders associated with inopportune invasion of  
 CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
 CC of prematurity, macular degeneration, corneal graft rejection,  
 CC retrolental fibroplasia, rubeosis, capillary proliferation in  
 CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
 CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints and wound  
 CC granulation. In addition, the peptides can be used as birth control  
 CC agents  
 XX  
 SQ Sequence 184 AA;  
 Query Match 66.1%; Score 39; DB 4; Length 184;  
 Best Local Similarity 81.8%; Pred. No. 42;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLPI 11  
 | | | | | | | |  
 Db 164 SCHNSYIVLCI 174

RESULT 13  
 ABG31793  
 ID ABG31793 standard; protein; 184 AA.  
 XX  
 AC ABG31793;  
 XX  
 DT 05-DEC-2002 (first entry)  
 XX  
 DE Human endostatin polypeptide.  
 XX

KW Human; endostatin; tumour; cancer; metastasis; cytostatic;  
 XX antiangiogenic.

OS Homo sapiens.

XX WO200268457-A2.

PN 06-SEP-2002.

XX 27-FEB-2002; 2002WO-IT000119.

XX 27-FEB-2001; 2001IT-MI000394.

XX (UYMI-) UNIV MILANO.

XX Chillemi F, Vicentinie LMT, Francescato P;

XX WPI; 2002-69855/75.

XX New peptide useful for the preparation of medicaments with antiangiogenic  
 PT activity that may be used in treating tumors or metastases, comprises a  
 PT sequence corresponding to fragments of human endostatin.

XX Disclosure; Fig 1; 24pp; English.

XX The invention relates to peptide comprising 20-50 amino acids with  
 CC sequences corresponding to the human endostatin polypeptide sequence, its  
 CC salt or non-toxic derivative. The peptides are useful in the preparation  
 CC of medicaments with antiangiogenic activity which may be useful in  
 CC treating tumours or metastases. This sequence represents a human  
 CC endostatin polypeptide

XX Sequence 184 AA;

Query Match 66.1%; Score 39; DB 5; Length 184;

Best Local Similarity 81.8%; Pred. No. 42;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11

DB 164 SCHNSYIVLCI 174

RESULT 14

ABU64290

ID ABU64290 standard; protein; 184 AA.

XX AC ABU64290;

DT 11-MAR-2004 (first entry)

DE Murine endostatin protein.

XX Vector; rAAV; recombinant adeno-associated viral vector;  
 KW anti-angiogenesis; PDF; angiogenesis; eye disorder; blindness;  
 KW retinal degeneration; macular degeneration; neovascularisation;  
 KW ophthalmological.

OS Mus sp.

XX WO2003080648-A2.

XX 02-OCT-2003.

XX 20-MAR-2003; 2003WO-US008667.

XX 20-MAR-2002; 2002US-0366114P.

XX (UYFL) UNIV FLORIDA RES FOUND INC.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Hauswirth WW, Campochiaro PA, Berns KI;

XX

DR WPI; 2003-779243/73.

XX N-PSDB; AAL56259.

PT Novel adeno-associated viral vector comprising polynucleotide encoding  
 PT pigment epithelium-derived factor, useful for treating choroidal  
 PT neovascularization, blindness, loss of vision.

XX Claim 14; Page 38; Opp; English.

XX The present invention relates to an adeno-associated viral (AAV) vector  
 CC comprising a polynucleotide that comprises a nucleic acid segment that  
 CC encodes a choroidal or ocular neovascularisation inhibitory polypeptide  
 CC operably linked to a promoter that expresses the segment to produce the  
 CC polypeptide in a selected mammalian host cell. Such a vector is useful  
 CC for providing a choroidal or ocular neovascularisation inhibitory  
 CC polypeptide to a mammal, for use in the therapy of ocular  
 CC neovascularisation, choroidal neovascularisation, retinal  
 CC neovascularisation, age-related macular degeneration, visual impairment,  
 CC ocular dysfunction, loss of vision, retinopathy, or blindness in a human.  
 CC The present sequence is a protein shown in the exemplification of the  
 CC invention

SQ Sequence 184 AA;

Query Match 66.1%; Score 39; DB 7; Length 184;

Best Local Similarity 81.8%; Pred. No. 42;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11

DB 164 SCHNSYIVLCI 174

RESULT 15

AAB28398

ID AAB28398 standard; protein; 191 AA.

XX AC AAB28398;

DT 19-FEB-2001 (first entry)

DE Murine endostatin.

XX Murine; endostatin; cytostatic; antiproliferative;  
 KW vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;  
 KW cancer; vascularised solid tumour.

XX OS Mus sp.

XX WO200064946-A2.

XX 02-NOV-2000.

XX 28-APR-2000; 2000WO-US011367.

XX 28-APR-1999; 99US-0131432P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Thorpe PE, Brekken RA;

XX WPI; 2000-687317/67.

XX N-PSDB; AAC67777.

XX Immunogenic composition for the treatment and diagnosis of cancer  
 PT comprises an anti-VEGF (vascular endothelial growth factor) antibody  
 PT binding the same epitope as the monoclonal antibody ATCC PTA 1595.

XX Example 10; Page 290-291; 298pp; English.

XX The present invention relates to anti-Vascular Endothelial Growth Factor  
 CC (VEGF) antibodies that bind to the same epitope as the monoclonal  
 CC antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to

CC the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF  
 CC receptor VEGFR1. The present sequence is murine endostatin. Endostatin  
 CC may be conjugated onto the anti-VEGF antibodies of the present invention.  
 CC The anti-VEGF antibodies of the present invention are useful for the  
 CC treatment and diagnosis of cancer, especially vascularised solid tumours  
 XX  
 SQ Sequence 191 AA;

Query Match 66.1%; Score 39; DB 3; Length 191;  
 Best Local Similarity 81.8%; Pred. No. 43;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
 | | | | | | | | | |  
 Db 171 SCHNSYIVLCI 181

Search completed: March 10, 2005, 11:06:50  
 Job time : 57.8684 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 10:59:54 ; Search time 56.3684 Seconds  
(without alignments)  
127.183 Million cell updates/sec

Title: US-09-766-412-31

Perfect score: 68

Sequence: 1 VPVNLKDEVLS 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_spot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	160	2 Q9CRT2	Q9crt2 mus musculus
2	68	100.0	171	2 Q9WUW5	Q9wuw5 rattus norv
3	68	100.0	226	2 Q9QZD2	Q9qzd2 rattus norv
4	68	100.0	1140	2 Q61434	Q61434 mus musculus
5	68	100.0	1315	2 Q6NZK9	Q6nzk9 mus musculus
6	68	100.0	1315	2 Q6F1Y4	Q6f1y4 mus musculus
7	68	100.0	1774	1 CALH_MOUSE	P39061 mus musculus
8	59	86.8	187	2 Q8WXI5	Q8wx15 homo sapien
9	59	86.8	261	2 Q8NG19	Q8ng19 homo sapien
10	59	86.8	816	2 Q8N4S4	Q8n4s4 homo sapien
11	59	86.8	1336	2 Q6RZ41	Q6rz41 homo sapien
12	59	86.8	1516	1 CALH_HUMAN	P39060 homo sapien
13	59	86.8	1516	2 Q6RZ39	Q6rz39 homo sapien
14	59	86.8	1751	2 Q6RZ40	Q6rz40 homo sapien
15	50	73.5	1344	2 Q93419	Q93419 gallus gall
16	48	70.6	102	2 Q96T70	Q96t70 homo sapien
17	48	70.6	361	2 Q8AWC6	Q8awc6 brachydanto
18	44	64.7	320	2 Q83V30	Q83v30 pseudomonas
19	43	63.2	327	2 Q9RZF8	Q9rzf8 caenorhabdi
20	42	61.8	582	2 Q6FTP7	Q6ftp7 candida gla
21	42	61.8	1301	2 Q6P3M9	Q6p3m9 xenopus tro
22	42	61.8	1307	2 Q8JFF7	Q8jff7 xenopus lae
23	42	61.8	1310	2 Q6DPR4	Q6dpr4 xenopus tro
24	42	61.8	1315	2 Q8QHL9	Q8qhl9 xenopus lae
25	41.5	61.0	778	2 Q7RZ49	Q7rz49 neurospora
26	41	60.3	239	2 Q8NSN5	Q8nsn5 corynebacte
27	41	60.3	310	2 Q7UP18	Q7up18 rhodospirell
28	41	60.3	323	2 Q810Y8	Q810y8 caenorhabdi
29	41	60.3	393	2 Q6Z165	Q6z165 oryza sativ
30	41	60.3	640	2 Q9NA34	Q9na34 caenorhabdi
31	41	60.3	795	1 P5CS_HUMAN	P54886 h delta 1-p

32 41 60.3 795 1 P5CS\_MOUSE  
33 41 60.3 795 2 Q8BGM2  
34 41 60.3 825 2 Q8FR36  
35 41 60.3 1334 2 Q898I5  
36 40 58.8 248 2 Q8YXX9  
37 40 58.8 270 2 Q8RPL8  
38 40 58.8 274 2 Q8TWES  
39 40 58.8 399 2 Q9ZTA2  
40 40 58.8 432 2 Q944Z9  
41 40 58.8 432 2 Q945B5  
42 40 58.8 468 2 P96758  
43 40 58.8 468 2 Q8G888  
44 40 58.8 468 2 Q8G8Y1  
45 40 58.8 468 2 Q8GGU9

Q9z110 m delta 1-p  
Q8bgm2 mus musculus  
Q8fr36 corynebacte  
Q898i5 clostridium  
Q8yxx9 rhyzopertha  
Q8rpl8 ehrlichia c  
Q8twes methanopyru  
Q9zta2 arabidopsis  
Q944z9 arabidopsis  
Q945b5 arabidopsis  
P96758 ehrlichia c  
Q8g888 ehrlichia c  
Q8g8y1 ehrlichia c  
Q8gg9u ehrlichia c

#### ALIGNMENTS

RESULT 1  
Q9CRT2 PRELIMINARY; PRT; 160 AA.  
AC Q9CRT2;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched  
DE library, clone:3200001M10 product:procollagen, type XVIII, alpha 1,  
DE full insert sequence. (Fragment).  
GN Name:Coll1a1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning."  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA The FANTOM Consortium,  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RL "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RN Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes."  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RL sequencing pipeline with 384 multicapillary sequencer.",  
 RN Genom. Res. 10:1757-1771(2000).  
 RP [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Konda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.,  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK014292; BAB29249.1; -.  
 DR HSSP; P39061; IDY1.  
 DR MGD; MGI:89451; Coll18a1.  
 DR GO; GO:0005604; C:basement membrane; IDA.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0001525; P:angiogenesis; IMP.  
 DR GO; GO:0001886; P:endothelial cell morphogenesis; IDA.  
 DR GO; GO:0043065; P:positive regulation of apoptosis; IDA.  
 DR GO; GO:0030335; P:positive regulation of cell migration; IDA.  
 DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.  
 DR InterPro; IPR010515; Endostatin.  
 DR Pfam; PF06482; Endostatin; 1.  
 KW Collagen.  
 FT NON\_TER 1  
 SQ SEQUENCE 160 AA; 17725 MW; 60F853D777C375D2 CRC64;  
 Query Match 100.0%; Score 68; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 0.00043;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 45 VPVNLKDEVLSPS 58  
 |||||  
 RESULT 2  
 Q9WUW5 PRELIMINARY; PRT; 171 AA.  
 AC Q9WUW5;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Collagen type XVIII, alpha (I) chain (Fragment).  
 GN Name=coll18a1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=21476596; PubMed=11592600; DOI=10.1016/S0168-8278(01)00134-9;  
 RA Jia J.D., Bauer M., Sedlacek N., Herbst H., Ruehl M., Hahn E.G.,  
 RA "Modulation of collagen XVIII/endostatin expression in lobular and  
 RT biliary rat liver fibrogenesis";  
 RL J. Hepatol. 35:386-391(2001).  
 DR EMBL; AJ236873; CAB44263.1; -.  
 DR HSSP; P39061; IDY1.  
 DR InterPro; IPR010515; Endostatin.  
 DR Pfam; PF06482; Endostatin; 1.  
 KW Collagen.

FT NON\_TER 1  
 FT NON\_TER 171  
 SQ SEQUENCE 171 AA; 18933 MW; 81BE2BE3FC2C8E72 CRC64;  
 Query Match 100.0%; Score 68; DB 2; Length 171;  
 Best Local Similarity 100.0%; Pred. No. 0.00046;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 61 VPVNLKDEVLSPS 74  
 |||||  
 RESULT 3  
 Q9QZD2 PRELIMINARY; PRT; 226 AA.  
 AC Q9QZD2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Collagen XVIII (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=2027226; PubMed=10766159;  
 RA Perletti G., Concarì P., Giardini R., Mairas E., Piccinini F.,  
 RA Folkman J., Chen L.;  
 RT "Antitumor activity of endostatin against carcinogen-induced rat  
 RL primary mammary tumors";  
 RL Cancer Res. 60:1793-1796(2000).  
 DR EMBL; AF189709; AAF00975.1; -.  
 DR HSSP; P39061; IDY1.  
 DR InterPro; IPR010515; Endostatin.  
 DR Pfam; PF06482; Endostatin; 1.  
 KW Collagen.  
 FT NON\_TER 1  
 SQ SEQUENCE 226 AA; 25350 MW; 38B83C0486C0E949 CRC64;  
 Query Match 100.0%; Score 68; DB 2; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 0.00063;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 111 VPVNLKDEVLSPS 124  
 |||||  
 RESULT 4  
 Q61434 PRELIMINARY; PRT; 1140 AA.  
 AC Q61434;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Collagen (Fragment).  
 GN Name=Coll15a1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Abe N., Muragaki Y., Yoshioka H., Inoue H., Ninomiya Y.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D17546; BAA04483.1; -.  
 DR FIR; B56101; B56101.  
 DR HSSP; P39061; IDY1.  
 DR MGD; MGI:88449; Coll15a1.  
 DR GO; GO:0005737; C:cytoplasm; IEA.



```
DR GO: 0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR010515; Endostatin.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD000007; Clg_helix; 3.
KW Collagen.
FT NON_TER
SQ SEQUENCE 1140 AA; 115156 MW; 8B0C7E6862B3BDFF CRC64;
Query Match 100.0%; Score 68; DB 2; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPVNLKDEVLSPS 14
DB 1025 VPVNLKDEVLSPS 1038
|||||
RESULT 5
Q6NZK9 PRELIMINARY; PRT; 1315 AA.
ID Q6NZK9
AC Q6NZK9
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Procollagen, type XVIII, alpha 1.
GN Name=Coll18a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066080; AAH66080.1; -.
DR GO: 0005604; C:basement membrane; IDA.
DR GO: 0005615; C:extracellular space; TAS.
DR GO: 0001525; P:angiogenesis; IMP.
DR GO: 0001886; P:endothelial cell morphogenesis; IDA.
DR GO: 0003035; P:positive regulation of apoptosis; IDA.
DR GO: 0003035; P:positive regulation of cell migration; IDA.
DR GO: 0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR003129; TSP_N.
```

```
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF01391; Collagen; 7.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD000007; Clg_helix; 2.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1315 AA; 134245 MW; 15C862C1E385F03A CRC64;
Query Match 100.0%; Score 68; DB 2; Length 1315;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPVNLKDEVLSPS 14
DB 1200 VPVNLKDEVLSPS 1213
|||||
RESULT 6
Q6PIV4 PRELIMINARY; PRT; 1315 AA.
ID Q6PIV4
AC Q6PIV4
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Procollagen, type XVIII, alpha 1.
GN Name=Coll18a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064817; AAH64817.1; -.
DR GO: 0005604; C:basement membrane; IDA.
DR GO: 0005615; C:extracellular space; TAS.
DR GO: 0001525; P:angiogenesis; IMP.
DR GO: 0001886; P:endothelial cell morphogenesis; IDA.
DR GO: 0003035; P:positive regulation of apoptosis; IDA.
DR GO: 0003035; P:positive regulation of cell migration; IDA.
DR GO: 0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR003129; TSP_N.
```

DR Pfam: PF01391; Collagen: 7.  
DR Pfam: PF06482; Endostatin; 1.  
DR ProDom: PD000007; C1g\_helix; 2.  
DR SMART: SM00210; TSPN; 1.  
SQ COLLAGEN.  
KW SEQUENCE 1315 AA; 134203 MW; B8215602ACE7AD1F CRC64;  
  
Query Match 100.0%; Score 68; DB 2; Length 1315;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 VPVNLKDEVLSPS 14  
||| |||||  
Db 1200 VPVNLKDEVLSPS 1213  
  
RESULT 7  
CA1H MOUSE  
ID CA1H MOUSE STANDARD; PRT: 1774 AA.  
AC P39061; Q60672; Q61437; Q62001; Q62002; Q9JK63;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].  
GN Name=Col18a1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RC STRAIN=BALB/c; TISSUE=Liver;  
RX MEDLINE=94245707; PubMed=8188673;  
RA Rehn M.V., Hintikka E., Pihlajaniemi T.;  
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,  
RT partial structure of the corresponding gene, and comparison of the  
RT alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen  
RT chain.";  
RL J. Biol. Chem. 269:13929-13935 (1994).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE=96435922; PubMed=8838808; DOI=10.1006/geno.1996.0139;  
RA Rehn M., Hintikka E., Pihlajaniemi T.;  
RT "Characterization of the mouse gene for the alpha-1 chain of type  
RT XVIII collagen (COL18A1) reveals that the three variant N-terminal  
RT polypeptide forms are transcribed from two widely separated  
RT promoters.";  
RL Genomics 32:436-446 (1996).  
RN [3]  
RP SEQUENCE OF 1-1387 FROM N.A. (ISOFORM 3).  
RX MEDLINE=94240112; PubMed=8183894;  
RA Rehn M.V., Pihlajaniemi T.;  
RT "Alpha 1(XVIII), a collagen chain with frequent interruptions in the  
RT collagenous sequence, a distinct tissue distribution, and homology  
RT with type XV collagen.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238 (1994).  
RN [4]  
RP SEQUENCE OF 1-562 FROM N.A. (ISOFORMS 1 AND 3), AND TISSUE  
RP SPECIFICITY.  
RX MEDLINE=95181468; PubMed=7876242; DOI=10.1074/jbc.270.9.4705;  
RA Rehn M., Pihlajaniemi T.;  
RT "Identification of three N-terminal ends of type XVIII collagen chains  
RT and tissue-specific differences in the expression of the corresponding  
RT transcripts. The longest form contains a novel motif homologous to rat  
RT and Drosophila frizzled proteins.";  
RL J. Biol. Chem. 270:4705-4711 (1995).  
RN [5]  
RP SEQUENCE OF 487-1774 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=94240111; PubMed=8183893;  
RA Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Ooshima A., Olsen B.R.;  
RT "Isolation and sequencing of cDNAs for proteins with multiple domains  
RT of Gly-Xaa-Yaa repeats identify a distinct family of collagenous  
RT proteins.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233 (1994).  
RN [6]  
RP SEQUENCE OF 1591-1774 FROM N.A.  
RX MEDLINE=21217748; PubMed=11321448;  
RA Jia S., Zhu F., Li H., He F., Xiu R.-J.;  
RT "Anticancer treatment of endostatin gene therapy by targeting tumor  
RT neovasculature in C57/BL mice.";  
RL Clin. Hemorheol. Microcirc. 23:251-257 (2000).  
RN [7]  
RP CHARACTERIZATION OF ENDOSTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE=97160848; PubMed=9008168; DOI=10.1016/S0092-8674(00)81848-6;  
RA O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,  
RA Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;  
RT "Endostatin: an endogenous inhibitor of angiogenesis and tumor  
RT growth.";  
RL Cell 88:277-285 (1997).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.  
RX MEDLINE=98169382; PubMed=9501087; DOI=10.1093/emboj/17.6.1656;  
RA Hohenester E., Sasaki T., Olsen B.R., Timpl R.;  
RT "Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A  
RT resolution.";  
RL EMBO J. 17:1656-1664 (1998).  
CC -!- FUNCTION: Endostatin potentially inhibits endothelial cell  
CC proliferation and angiogenesis. May inhibit angiogenesis by  
CC binding to the heparan sulfate proteoglycans involved in growth  
CC factor signaling.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative promoter;  
CC Comment=2 isoforms, 1 (shown here) and 3, are produced by use of  
CC alternative promoters;  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=1; Synonyms=NCl-764;  
CC IsoId=P39061-3; Sequences=Displayed;  
CC Name=2; Synonyms=Long, NCl-517;  
CC IsoId=P39061-1; Sequences=VSP 008303;  
CC Note=Produced by alternative splicing of isoform 1;  
CC Name=3; Synonyms=Short, NCl-301; Sequences=VSP 001157, VSP 001158;  
CC -!- TISSUE SPECIFICITY: Expressed in liver, kidney, lung, skeletal  
CC muscle and testis.  
CC -!- PTM: Prolines at the third position of the tripeptide repeating  
CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
CC -!- SIMILARITY: belongs to the fibril-associated collagens with  
CC interrupted helices (FACIT) family.  
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L16898; AAA37434.1; -  
CC EMBL; U03714; AAA20657.1; -  
CC EMBL; U03715; AAC52901.1; -  
CC EMBL; U34606; AAC52901.1; JOINED.  
CC EMBL; U34608; AAC52901.1; JOINED.  
CC EMBL; U34609; AAC52901.1; JOINED.  
CC EMBL; U34610; AAC52901.1; JOINED.  
CC EMBL; U34611; AAC52901.1; JOINED.  
CC EMBL; U34612; AAC52901.1; JOINED.  
CC EMBL; U34613; AAC52901.1; JOINED.  
CC EMBL; U03716; AAC52901.1; JOINED.  
CC EMBL; U03718; AAC52901.1; JOINED.  
CC EMBL; U03715; AAC52902.1; -  
CC EMBL; U34607; AAC52902.1; JOINED.  
CC EMBL; U34608; AAC52902.1; JOINED.  
CC EMBL; U34609; AAC52902.1; JOINED.

Query Match 100.0%; Score 68; DB 1; Length 1774;  
Best Local Similarity 100.0%; Pred. No. 0.0059;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 VPVNLKDEVLSPS 14  
|||||  
Db 1659 VPVNLKDEVLSPS 1672

RESULT 8  
Q8WX15 PRELIMINARY; PRT; 187 AA.  
AC Q8WX15;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Collagen XVIII (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RX SEQUENCE FROM N.A.  
RX MEDLINE=21409408; PubMed=11517600;  
RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.;  
RT "Inhibition effect in vitro of purified endostatin expressed in Pichia pastoris.";  
RL Sheng Wu Gong Cheng Xue Bao 17:278-282(2001).  
RN [1]  
RP SEQUENCE FROM N.A.  
RP MEDLINE=21409408; PubMed=11517600;  
RA Feng Y., Wu Y., Zhu X., Liu C.X., Ma Q.J.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF416592; AAL37720.1; -;  
DR HSSP; P39060; 1BNL.  
KW Pfam; PF06482; Endostatin; 1.  
KW Collagen.  
FT NON\_TER 1 1  
SQ SEQUENCE 187 AA; 20448 MW; 72B1047D85838CD3 CRC64;

Query Match 86.8%; Score 59; DB 2; Length 187;  
Best Local Similarity 85.7%; Pred. No. 0.022;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps

QY 1 VPVNLKDEVLSPS 14  
|||||  
Db 73 VPVNLKDEVLSPS 86

RESULT 9  
Q8NG19 PRELIMINARY; PRT; 261 AA.  
AC Q8NG19;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Multi-functional protein MFP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dou D.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF282883; AAM52249.1; -;  
DR HSSP; P39060; 1BNL.  
DR InterPro; IPR010515; Endostatin.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF06482; Endostatin; 1.  
DR ProDom; PD000395; Kringle; 1.  
KW Kringle.  
SQ SEQUENCE 261 AA; 26745 MW; CA60C920AF3E90E5 CRC64;

Query Match 100.0%; Score 68; DB 1; Length 1774;  
Best Local Similarity 100.0%; Pred. No. 0.0059;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 VPVNLKDEVLSPS 14  
|||||  
Db 1659 VPVNLKDEVLSPS 1672

RESULT 8  
Q8WX15 PRELIMINARY; PRT; 187 AA.  
AC Q8WX15;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Collagen XVIII (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RX SEQUENCE FROM N.A.  
RX MEDLINE=21409408; PubMed=11517600;  
RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.;  
RT "Inhibition effect in vitro of purified endostatin expressed in Pichia pastoris.";  
RL Sheng Wu Gong Cheng Xue Bao 17:278-282(2001).  
RN [1]  
RP SEQUENCE FROM N.A.  
RP MEDLINE=21409408; PubMed=11517600;  
RA Feng Y., Wu Y., Zhu X., Liu C.X., Ma Q.J.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF416592; AAL37720.1; -;  
DR HSSP; P39060; 1BNL.  
KW Pfam; PF06482; Endostatin; 1.  
KW Collagen.  
FT NON\_TER 1 1  
SQ SEQUENCE 187 AA; 20448 MW; 72B1047D85838CD3 CRC64;

Query Match 86.8%; Score 59; DB 2; Length 187;  
Best Local Similarity 85.7%; Pred. No. 0.022;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps

QY 1 VPVNLKDEVLSPS 14  
|||||  
Db 73 VPVNLKDEVLSPS 86

RESULT 9  
Q8NG19 PRELIMINARY; PRT; 261 AA.  
AC Q8NG19;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Multi-functional protein MFP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dou D.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF282883; AAM52249.1; -;  
DR HSSP; P39060; 1BNL.  
DR InterPro; IPR010515; Endostatin.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF06482; Endostatin; 1.  
DR ProDom; PD000395; Kringle; 1.  
KW Kringle.  
SQ SEQUENCE 261 AA; 26745 MW; CA60C920AF3E90E5 CRC64;

```

Query Match      86.8%; Score 59; DB 2; Length 261;
Best Local Similarity 85.7%; Pred. No. 0.032;
Matches 12; Conservative 1; Mismatches 0; Gaps 0; Indels 1;

QY 1 VP1VNLKDEVLSPS 14
DB 147 VP1VNLKDELLFPS 160

RESULT 10
Q8N4S4
ID Q8N4S4 PRELIMINARY; PRT; 816 AA.
AC Q8N4S4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to collagen, type XVIII, alpha 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033715; AAH33715.1; -.
DR HSSP; P39060; 1BNL.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR010515; Endostatin.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD000007; Clg_helix; 2.
KW Collagen.
FT NON TER
SQ SEQUENCE 816 AA; 82553 MW; 5D539B2946694F86 CRC64;

Query Match      86.8%; Score 59; DB 2; Length 816;
Best Local Similarity 85.7%; Pred. No. 0.11;
Matches 12; Conservative 1; Mismatches 0; Gaps 0; Indels 1;

QY 1 VP1VNLKDEVLSPS 14
DB 702 VP1VNLKDELLFPS 715

RESULT 11
Q6RZ41
ID Q6RZ41 PRELIMINARY; PRT; 1336 AA.
AC Q6RZ41;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type XVIII collagen short variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;
RA Elanah H., Snellman A., Rehn M., Autio-Harminen H., Pihlajaniemi T.;
RT "Characterization of the human type XVIII collagen gene and
RT proteolytic processing and tissue location of the variant containing a
RT frizzled motif.";
RL Matrix Biol. 22:427-442(2003).
DR EMBL; AY484967; AAR83298.1; JOINED.
DR EMBL; AY484969; AAR83298.1; JOINED.
DR EMBL; AY484970; AAR83298.1; JOINED.
DR EMBL; AY484971; AAR83298.1; -.

Query Match      86.8%; Score 59; DB 2; Length 1336;
Best Local Similarity 85.7%; Pred. No. 0.19;
Matches 12; Conservative 1; Mismatches 0; Gaps 0; Indels 1;

QY 1 VP1VNLKDEVLSPS 14
DB 1222 VP1VNLKDELLFPS 1235

RESULT 12
CA1H HUMAN
ID CA1H HUMAN STANDARD; PRT; 1516 AA.
AC P39060; Q9UK38; Q9Y6Q7; Q9Y6Q8;
DT 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].
GN Name=COL18A1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.; AND VARIANT ILE-841.
RX MEDLINE=98164096; PubMed=9503365; DOI=10.1016/S0945-053X(98)90003-8;
RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;
RT "Complete primary structure of two variant forms of human type XVIII
RT collagen and tissue-specific differences in the expression of the
RT corresponding transcripts.";
RL Matrix Biol. 16:319-328(1998).
RN 2
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN 3
RP SEQUENCE OF 834-1516 FROM N.A. AND VARIANT ARG-886.
RX MEDLINE=94245237; PubMed=8188291;
RA Oh S.P., Warman M.L., Seldin M.F., Cheng S., Knoll J.H., Timmons S.,
RA Olsen B.R.;
RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen
RT and localization of the alpha 1(XVIII) collagen gene to mouse
RT chromosome 10 and human chromosome 21.";
RL Genomics 19:494-499(1994).
RN 4

```



```

FT VARIANT 1437 1437 /FTid=VAR_018056.
FT D-->N (decreased activity for binding
FT laminin; increased risk of developing
FT prostate cancer; in compound
FT heterozygotes may cause Knobloch syndrome
FT when in combination with a frameshift/
FT truncating mutation).
FT /FTid=VAR_012709.

Query Match 86.8%; Score 59; DB 1; Length 1516;
Best Local Similarity 85.7%; Pred. No. 0.22;
Matches 12; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 1402 VPVNLKDEVLSPS 1415

RESULT 13
Q6R239 ID Q6R239 PRELIMINARY; PRT; 1516 AA.
AC Q6R239
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type XVII collagen middle variant.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;
RA Elamasa H., Snellman A., Rehn M., Autio-Harmanen H., Pihlajaniemi T.;
RT "Characterization of the human type XVIII collagen gene and
RT proteolytic processing and tissue location of the variant containing a
RT frizzled motif.";
RL Matrix Biol. 22:427-442(2003).
DR EMBL; AY484968; AAR83297.1; JOINED.
DR EMBL; AY484969; AAR83297.1; JOINED.
DR EMBL; AY484970; AAR83297.1; JOINED.
DR EMBL; AY484971; AAR83297.1; JOINED.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR010363; DUF959.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF01391; Collagen_6.
DR Pfam; PF06121; DUF959; 1.
DR Pfam; PF06482; Endostatin; 1.
DR Pfam; PF01392; Fz; 1.
DR ProDom; PD000007; Clg_helix; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS50038; FZ; 1.
KW Collagen.
SQ SEQUENCE 1516 AA; 153779 MW; 38D0F23D3DFD758A4 CRC64;

Query Match 86.8%; Score 59; DB 2; Length 1516;
Best Local Similarity 85.7%; Pred. No. 0.22;
Matches 12; Conservative 1; Mismatches 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 1402 VPVNLKDEVLSPS 1415

RESULT 14
Q6R240 ID Q6R240 PRELIMINARY; PRT; 1751 AA.
AC Q6R240

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Type XVII collagen long variant.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;
RA Elamasa H., Snellman A., Rehn M., Autio-Harmanen H., Pihlajaniemi T.;
RT "Characterization of the human type XVIII collagen gene and
RT proteolytic processing and tissue location of the variant containing a
RT frizzled motif.";
RL Matrix Biol. 22:427-442(2003).
DR EMBL; AY484968; AAR83296.1; JOINED.
DR EMBL; AY484969; AAR83296.1; JOINED.
DR EMBL; AY484970; AAR83296.1; JOINED.
DR EMBL; AY484971; AAR83296.1; JOINED.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; ConA_like Lec_gl.
DR InterPro; IPR010363; DUF959.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF01391; Collagen_6.
DR Pfam; PF06121; DUF959; 1.
DR Pfam; PF06482; Endostatin; 1.
DR Pfam; PF01392; Fz; 1.
DR ProDom; PD000007; Clg_helix; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS50038; FZ; 1.
KW Collagen.
SQ SEQUENCE 1751 AA; 177920 MW; 481F4F72176FA2A6 CRC64;

Query Match 86.8%; Score 59; DB 2; Length 1751;
Best Local Similarity 85.7%; Pred. No. 0.26;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 1637 VPVNLKDEVLSPS 1650

RESULT 15
Q93419 ID Q93419 PRELIMINARY; PRT; 1344 AA.
AC Q93419
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen XVIII precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98411346; PubMed=9738008; DOI=10.1074/jbc.273.39.25404;
RA Halfter W., Dong S., Schurer B., Cole G.J.;
RT "Collagen XVIII is a basement membrane heparan sulfate proteoglycan."
RL J. Biol. Chem. 273:25404-25412(1998).

```

[2]  
RN SEQUENCE FROM N.A.  
RP Halfter W., Dong S.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF083440; AAC33294.2; --  
DR HSP; F39060; IBNL.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0007155; P:cell adhesion; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008161; Cig\_helix.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR008985; ConA\_like\_lectin.  
DR InterPro; IPR010515; Endostatin.  
DR InterPro; IPR01791; Laminin\_G.  
DR InterPro; IPR003129; TSP\_N.  
DR Pfam; PF01391; Collagen\_8.  
DR Pfam; PF06482; Endostatin; 1.  
DR ProDom; PD000007; Cig\_helix; 3.  
DR SMART; SM00282; LamG; 1.  
DR SMART; SM00210; TSPN; 1.  
KW Collagen; Signal.  
FT SIGNAL 1 26 Potential.  
SQ SEQUENCE 1344 AA; 137402 MW; 7AA366E4FE940CCD CRC64;

Query Match 73.5%; Score 50; DB 2; Length 1344;  
Best Local Similarity 90.9%; Pred. No. 8.4;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VPVNLKDEVL 11  
|||||:||||  
Db 1229 VPVNLKDEVL 1239

Search completed: March 10, 2005, 11:09:31  
Job time : 57.3684 secs

**This Page Blank (uspto)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:00:49 ; Search time 12.1579 Seconds  
(without alignments)  
110.795 Million cell updates/sec

Title: US-09-766-412-31  
Perfect score: 68  
Sequence: 1 VP1VNLKDEVLSPS 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	1315	A56101	collagen alpha 1(X)
2	58	100.0	1774	B56101	collagen alpha 1(X)
3	59	86.8	684	A53019	collagen alpha 1(X)
4	43	63.2	327	T33540	hypothetical prote
5	40	58.8	399	T01387	oxidoreductase hom
6	40	58.8	548	JC6174	immunodominant 120
7	40	58.8	1388	A53317	collagen alpha 1(X)
8	39	57.4	187	E90390	hypothetical prote
9	39	57.4	264	D71724	hypothetical prote
10	39	57.4	266	A97723	hypothetical prote
11	39	57.4	887	T41543	DNA topoisomerase
12	38	55.9	334	T04198	hypothetical prote
13	38	55.9	485	E88114	protein F53C3.9 li
14	38	55.9	535	S68446	origin recognition
15	38	55.9	565	H82516	60kDa inner-membra
16	38	55.9	813	D70176	endopeptidase Ia (
17	38	55.9	1016	T25433	hypothetical prote
18	38	55.9	1222	S40977	hypothetical prote
19	37.5	55.1	425	H83652	seryl-tRNA synthet
20	37	54.4	107	AF1948	hypothetical prote
21	37	54.4	170	T15452	hypothetical prote
22	37	54.4	190	F70410	hypothetical prote
23	37	54.4	261	G69134	hypothetical prote
24	37	54.4	308	H85214	nodulin-26-like pr
25	37	54.4	319	S73159	hypothetical prote
26	37	54.4	337	T20947	hypothetical prote
27	37	54.4	364	A89004	protein T24A6.8 li
28	37	54.4	404	S42831	F40P12.2 protein -
29	37	54.4	441	T11715	probable acetylorn

30	37	54.4	442	2	T42091	acetylornithine tr
31	37	54.4	568	2	T39840	conserved hypothet
32	37	54.4	759	2	T16368	hypothetical prote
33	37	54.4	1043	2	T15191	hypothetical prote
34	36.5	53.7	425	2	S66043	serine-tRNA ligase
35	36	52.9	93	2	A95121	hypothetical prote
36	36	52.9	93	2	F97990	hypothetical prote
37	36	52.9	186	2	H97920	elongation factor
38	36	52.9	186	2	D95050	translation elonga
39	36	52.9	194	2	E75150	hypothetical prote
40	36	52.9	226	1	G64387	hypothetical prote
41	36	52.9	251	2	G75345	serine/threonine-p
42	36	52.9	341	2	JQ1189	CbHc protein - Cox
43	36	52.9	410	2	AB0077	probable aminotran
44	36	52.9	439	2	G64401	hypothetical prote
45	36	52.9	468	2	AE1942	two-component sens

## ALIGNMENTS

### RESULT 1

A56101 collagen alpha 1(XVIII) chain precursor, short splice form - mouse  
N:Contains: endostatin  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Oct-1995 #sequence revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: A56101; A58371; S72450; S65595  
R:Rehn, M.; Pihlajaniemi, T.  
J. Biol. Chem. 270, 4705-4711, 1995  
A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tiss  
tif homologous to rat and Drosophila frizzled proteins.  
A:Reference number: A56101; MUID:95181468; PMID:7876242  
A:Accession: A56101  
A:Molecule type: mRNA  
A:Residues: 1-103 <REHI>  
A:Cross-references: UNIPROT:P39061; GB:U11636; NID:G618427; PIDN:AAC52178.1; PID:G61842  
R:Rehn, M.; Pihlajaniemi, T.  
Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994  
A:Title: Alphas(XVIII), a collagen chain with frequent interruptions in the collagenous  
A:Reference number: A58371; MUID:94240112; PMID:8183894  
A:Accession: A58371  
A:Molecule type: mRNA  
A:Residues: 1-928 <REH2>  
A:Cross-references: GB:L16898; NID:9404754; PIDN:AAA37434.1; PID:G553894  
R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.  
submitted to the EMBL Data Library, August 1993  
A:Reference number: S72450  
A:Accession: S72450  
A:Molecule type: mRNA  
A:Residues: 28-687, 'L', 689-734, 'P', 736-751, 'R', 753-1315 <OHW>  
A:Cross-references: EMBL:L22545; NID:9348968; PIDN:AAA19787.1; PID:G511298  
R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.  
Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994  
A:Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xa  
A:Reference number: A58370; MUID:94240111; PMID:8183893  
A:Accession: S65595  
A:Molecule type: mRNA  
A:Residues: 28-1315 <OHS>  
A:Cross-references: EMBL:L22545  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
lated and subsequently O-glycosylated.  
C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per  
C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of u  
ay be useful in treating solid tumors.  
C:Genetics:  
A:Gene: MGI:Coll18a1  
A:Cross-references: MGI:71175  
A:Map position: 10:41.0  
C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:24-235/Region: thrombospondin amino-terminal homologous  
F:26-1315/Product: collagen alpha 1(XVIII) chain, short splice form #status predicted <

F:327-353/Domain: collagenous #status predicted <CO1>  
 F:364-437/Domain: collagenous #status predicted <CO2>  
 F:462-583/Domain: collagenous #status predicted <CO3>  
 F:607-689/Domain: collagenous #status predicted <CO4>  
 F:704-745/Domain: collagenous #status predicted <CO5>  
 F:759-831/Domain: collagenous #status predicted <CO6>  
 F:842-874/Domain: collagenous #status predicted <CO7>  
 F:882-910/Domain: collagenous #status predicted <CO8>  
 F:892-894/Region: cell attachment (R-G-D) motif  
 F:918-969/Domain: collagenous #status predicted <CO9>  
 F:983-1000/Domain: collagenous #status predicted <CO10>  
 F:1132-1315/Product: endostatin #status predicted <EST>  
 F:1139-1315/Region: multiplexin collagen carboxyl-terminal homologous  
 F:1267-488/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:172-228/Disulfide bonds: #status predicted  
 F:240,245,1257/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F:451,454,594/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 68; DB 2; Length 1315;  
 Best Local Similarity 100.0%; Pred. No. 0.00078; Mismatches 0; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VP1VNLKDEVLSPS 14  
 |||||  
 Db 1200 VP1VNLKDEVLSPS 1213

RESULT 2

B56101  
 N:Contains: collagen alpha 1(XVIII) chain precursor, long splice form - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 03-Oct-1995 #sequence revision 08-May-1998 #text\_change 09-Jul-2004  
 C:Accession: B56101; C56101; S72450; S65595; PN0675; A54072; A58816  
 R:Rehn, M.; Pihlajaniemi, T.  
 J. Biol. Chem. 270, 4705-4711, 1995

A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tiss  
 tif homologous co-rat and drosophila frizzled proteins.  
 A:Reference number: A56101; MUID:95181468; PMID:7876242  
 A:Accession: B56101  
 A:Molecule type: mRNA  
 A:Residues: 1-562-<REH1>  
 A:Cross-references: UNIPROT:O61434; GB:U11637; NID:G618429; PID:G618430  
 A:Experimental source: splice form clone P817.24  
 A:Accession: C56101  
 A:Molecule type: mRNA  
 A:Residues: 1-239,487-562 <REH2>  
 A:Cross-references: GB:U11637; NID:G618429  
 A:Experimental source: splice form clones PE8.1, PE19, PE15.2  
 R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.  
 submitted to the EMBL Data Library, August 1993  
 A:Reference number: S72450  
 A:Accession: S72450  
 A:Molecule type: mRNA  
 A:Residues: 487-1146, 'L', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'P', 1524-16  
 A:Cross-references: EMBL:U22545; NID:G348968; PID:AAA19787.1; PID:G511298  
 R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994

A:Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa  
 A:Reference number: A58370; MUID:94240111; PMID:8183893  
 A:Accession: S65595  
 A:Molecule type: mRNA  
 A:Residues: 487-1512, 'L', 1514-1522, 'F', 1524-1683, 'V', 1685-1774 <OH2>  
 A:Cross-references: EMBL:U22545  
 R:Abé, N.; Muragaki, Y.; Yoshioka, H.; Inoue, H.; Ninomiya, Y.  
 Biochem. Biophys. Res. Commun. 196, 576-582, 1993

A:Title: Identification of a novel collagen chain represented by extensive interruptions  
 A:Reference number: PN0675; MUID:94059075; PMID:8240330  
 A:Accession: PN0675  
 A:Molecule type: mRNA  
 A:Residues: 635-1774 <ABE>  
 R:Rehn, M.; Hintikka, E.; Pihlajaniemi, T.  
 J. Biol. Chem. 269, 13929-13935, 1994

A:Title: Primary structure of the alpha1 chain of mouse type XVIII collagen, partial seq  
 collagen chain.  
 A:Reference number: A54072; MUID:94245707; PMID:8188673  
 A:Accession: A54072  
 A:Molecule type: DNA; mRNA  
 A:Residues: 1293-1403, 'R', 1405-1774 <REH3>  
 A:Cross-references: GB:U03714; NID:G487733; PIDN:AAA20657.1; PID:G487734  
 R:O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W.S.; Flynn, E.; Bi  
 Cell 88, 277-285, 1997

A:Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.  
 A:Reference number: A58816; MUID:97160848; PMID:9008168  
 A:Accession: A58816  
 A:Molecule type: protein  
 A:Residues: 1591-1610 <ORE>  
 A:Experimental source: hemangioendothelium cells  
 A:Note: Inhibits endothelial cell proliferation  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 lated and subsequently O-glycosylated.  
 C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per  
 C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of u  
 ay be useful in treating solid tumors.

C:Genes: MGI:Coll18a1  
 A:Cross-references: MGI:71175  
 A:Map position: 10:41.0  
 A:Introns: 1295/3; 1310/1; 1331/1; 1345/3; 1388/3; 1437/1; 1461/3; 1505/3; 1516/3; 1599/  
 A:Note: the list of introns is incomplete  
 C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly  
 F:1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, long splice form #status pr  
 F:1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, medium splice form #;  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:786-812/Domain: collagenous #status predicted <CO01>  
 F:823-896/Domain: collagenous #status predicted <CO02>  
 F:921-1042/Domain: collagenous #status predicted <CO03>  
 F:1066-1148/Domain: collagenous #status predicted <CO04>  
 F:1163-1204/Domain: collagenous #status predicted <CO05>  
 F:1218-1290/Domain: collagenous #status predicted <CO06>  
 F:1301-1333/Domain: collagenous #status predicted <CO07>  
 F:1346-1369/Domain: collagenous #status predicted <CO08>  
 F:1351-1353/Region: cell attachment (R-G-D) motif  
 F:1437-1428/Domain: collagenous #status predicted <CO09>  
 F:1442-1459/Domain: collagenous #status predicted <CO10>  
 F:1591-1774/Product: endostatin #status predicted <EST>  
 F:1598-1774/Region: multiplexin collagen carboxyl-terminal homologous  
 F:354,361,947/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:699,704,1716/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F:910,913,1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 68; DB 2; Length 1774;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VP1VNLKDEVLSPS 14  
 |||||  
 Db 1659 VP1VNLKDEVLSPS 1672

RESULT 3

A53019  
 collagen alpha 1(XVIII) chain - human (fragment)  
 N:Contains: endostatin  
 C:Species: Homo sapiens (man)  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 15-Sep-2003  
 C:Accession: A53019  
 R:Oh, S.P.; Warman, M.L.; Seldin, M.F.; Cheng, S.D.; Knoll, J.H.M.; Timmons, S.; Olsen,  
 Genomics 19, 494-499, 1994

A:Title: Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localiz  
 A:Reference number: A53019; MUID:94245237; PMID:8188291  
 A:Accession: A53019  
 A:Molecule type: mRNA  
 A:Residues: 1-684 <OHA>  
 A:Cross-references: GB:L22548; NID:G349908; PIDN:AAA51864.1; PID:G562794  
 A:Note: the cited accession number, L25548, is not in Genbank release 103

A>Note: in the authors' translation, 482-Gly is not shown, residues 483-490 are shifted  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 lated and subsequently O-glycosylated.

C:Comment: Different splice forms of collagen alpha 1(XVIII) may be involved in perivas  
 C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un  
 ay be useful in treating solid tumors.

C:Genetics:

A:Gene: GDB:COL18A1

A:Cross-references: GDB:138752; OMIM:120328

A:Map position: 21q22.3-21q22.3

C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc  
 F1-684/Product: collagen alpha 1(XVIII) chain (fragment) #status predicted <CO4>

F1-55/Domain: collagenous (fragment) #status predicted <CO5>

F174-115/Domain: collagenous #status predicted <CO6>

F129-201/Domain: collagenous #status predicted <CO7>

F212-244/Domain: collagenous #status predicted <CO8>

F257-278/Domain: collagenous #status predicted <CO9>

F262-264/Region: cell attachment (R-G-D) motif

F286-340/Domain: collagenous #status predicted <CO9>

F354-371/Domain: collagenous #status predicted <CO10>

F502-684/Product: endostatin #status predicted <EST>

F509-684/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 86.8%; Score 59; DB 2; Length 684;

Best Local Similarity 85.7%; Pred. No. 0.016;

Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14

Db 570 VPVNLKDELLPFS 583

RESULT 4

T33540

hypothetical protein F58E1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T33540

R:Wameley, P.; Twyman, B.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F58E1.

A:Reference number: Z21365

A:Accession: T33540

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-327 <WAM>

A:Cross-references: UNIPROT:Q9TZF8; EMBL:AF098995; PIDN:AAC67479.1; GSPDB:GN00020; CESP:

A:Experimental source: strain Bristol N2; clone F58E1

C:Genetics:

A:Gene: CESP:F58E1.1

A:Map position: 2

A:Introns: 11/3

C:Superfamily: Caenorhabditis elegans ZK1236.4 protein

Query Match 63.2%; Score 43; DB 2; Length 327;

Best Local Similarity 64.3%; Pred. No. 5.6;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14

Db 185 VPVINYCEVYSPS 198

RESULT 5

T01387

oxidoreductase homolog T419.6 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C:Accession: T01387

R:Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.; Habermann, K.; Schutz, K.; Hua

submitted to the EMBL Data Library, May 1998

A:Description: Genomic sequence of BAC T419 from Arabidopsis thaliana, chromosome IV, ne

A:Reference number: Z14314

A:Accession: T01387

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-399 <PAR>

A:Cross-references: UNIPROT:Q9ZTA2; EMBL:AF069442; NID:G3242970; PID:G3924598

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Introns: 123/2; 167/3; 313/3

A:Note: T419.6

C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 58.8%; Score 40; DB 2; Length 399;

Best Local Similarity 46.2%; Pred. No. 25;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSP 13

Db 8 LPLINLADKTLEP 20

RESULT 6

JC6174

immunodominant 120K protein - Ehrlichia chaffeensis

C:Species: Ehrlichia chaffeensis

C>Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 09-Jul-2004

C:Accession: JC6174

R:Yu, X.; Crocquet-Valdes, P.; Walker, D.H.

Gene 184, 149-154, 1997

A:Title: Cloning and sequencing of the gene for a 120-kDa immunodominant protein of Ehr

A:Reference number: JC6174; MUID:97183655; PMID:9031621

A:Accession: JC6174

A:Molecule type: DNA

A:Residues: 1-548 <YUA>

A:Cross-references: UNIPROT:Q46976; GB:U9426; NID:G1864025; PIDN:AAC44861.1; PID:G1864

C:Comment: This protein is an immunodominant antigen of Ehrlichia chaffeensis which rea

athogenicity of Ehrlichia chaffeensis and stimulating protective immunity in humans.

Query Match 58.8%; Score 40; DB 2; Length 548;

Best Local Similarity 50.0%; Pred. No. 37;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14

Db 412 IPVWVEKDEMPAPS 425

RESULT 7

A53317

collagen alpha 1(XV) chain precursor - human

N:Alternate names: procollagen alpha 1(XV) chain

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A53317; A53146; S28778

R:Kivirikko, S.; Heinegaeki, P.; Rehn, M.; Honkanen, N.; Myers, J.C.; Pihlajaniemi, T.

J. Biol. Chem. 269, 4773-4779, 1994

A:Title: Primary structure of the alpha1 chain of human type XV collagen and exon-intro

A:Reference number: A53317; MUID:94148920; PMID:8106446

A:Accession: A53317

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1388 <KIV>

A:Cross-references: UNIPROT:P39059; GB:I25280

A:Note: nucleotide sequence and conceptual translation not complete

R:Muragaki, Y.; Abe, N.; Ninomiya, Y.; Olsen, B.R.; Ooshima, A.

J. Biol. Chem. 269, 4042-4046, 1994

A:Title: The human alpha1(XV) collagen chain contains a large amino-terminal non-triple

A:Reference number: A53146; MUID:94140817; PMID:8307960

A:Accession: A53146

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-9,'S',11-48,'V',50-94,'A',96-149,'A',151-203,'V',205-408,'A',410-569 <MUR

A:Cross-references: GB:D21230; NID:9415605; PIDN:BA04762.1; PID:di005294; PID:G460703

R.Myers, J.C.; Kivirikko, S.; Gordon, M.K.; Dion, A.S.; Pihlajaniemi, T.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 10144-10148, 1992  
 A>Title: Identification of a previously unknown human collagen chain, alpha1(XV), charac  
 A:Reference number: S28778; MUID:93066196; PMID:1279671  
 A:Accession: S28778  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 544-640, 'P', 642-811, 'P', 813-1252 <MYE>  
 C:Genetics:  
 A:Gene: GDB:COL15A1  
 A:Cross-references: GDB:132578; OMIM:120325  
 A:Map position: 9q21-9q22  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-1388/Product: collagen alpha 1(XV) chain #status predicted <WAT>  
 F:1216-1388/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 58.8%; Score 40; DB 2; Length 1388;  
 Best Local Similarity 72.7%; Pred. No. 1.1e+02;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVL 11  
 :|||||:|  
 DB 1273 LPVNLKGQVL 1283

RESULT 8  
 E90390  
 Hypothetical protein SSO2210 [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C:Accession: E90390  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, E.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: E90390  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-187 <KUR>  
 A:Cross-references: UNIPROT:Q97WK6; GB:AE006641; NID:gl3815510; PIDN:AAK42380.1; GSPDB:G  
 C:Genetics:  
 A:Gene: SSO2210

Query Match 57.4%; Score 39; DB 2; Length 187;  
 Best Local Similarity 63.6%; Pred. No. 15;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PIVNLKDEVLS 12  
 ||:|||||  
 DB 56 PIVDLKDDILS 66

RESULT 9  
 D71724  
 Hypothetical protein RP143 - Rickettsia prowazekii  
 C:Species: Rickettsia prowazekii  
 C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 09-Jul-2004  
 C:Accession: D71724  
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U  
 Nature 396, 133-140, 1998  
 A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
 A:Reference number: A7630; MUID:99039499; PMID:9823893  
 A:Accession: D71724  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-264 <AND>  
 A:Cross-references: UNIPROT:Q9ZE19; GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA1461  
 A:Experimental source: strain Madrid E  
 C:Genetics:  
 A:Gene: RP143

Query Match 57.4%; Score 39; DB 2; Length 264;  
 Best Local Similarity 63.6%; Pred. No. 23;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VNLKDEVLS 14  
 |:|||||  
 DB 142 VDSKDEIIISPS 152

RESULT 10  
 A97723  
 Hypothetical protein RC0185 [imported] - Rickettsia conorii (strain Malish 7)  
 C:Species: Rickettsia conorii  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: A97723  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R  
 Science 293, 2093-2098, 2001  
 A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: A97723  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-266 <KUR>  
 A:Cross-references: UNIPROT:Q92J84; GB:AE006914; PIDN:AAI02723.1; PID:gl5619234; GSPDB:G  
 C:Genetics:  
 A:Gene: RC0185

Query Match 57.4%; Score 39; DB 2; Length 266;  
 Best Local Similarity 63.6%; Pred. No. 24;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 VNLKDEVLS 14  
 |:|||||  
 DB 142 VDSKDEIIISPS 152

RESULT 11  
 S41543  
 DNA topoisomerase (EC 5.99.1.12) - Bacillus anthracis virulence plasmid pXO1  
 C:Species: Bacillus anthracis  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S41543; S78064; E59108  
 R:Fouet, A.; Sirard, J.C.; Mock, M.  
 Mol. Microbiol. 11, 471-479, 1994  
 A>Title: Bacillus anthracis pXO1 virulence plasmid encodes a type 1 DNA topoisomerase.  
 A:Reference number: S41543; MUID:94203050; PMID:8152371  
 A:Accession: S41543  
 A:Molecule type: DNA  
 A:Residues: 1-887 <FOU>  
 A:Cross-references: UNIPROT:P40114  
 A:Experimental source: strain 7702  
 R:Fouet, A.; Sirard, J.C.; Mock, M.  
 submitted to the EMBL Data Library, May 1994  
 A:Description: Bacillus anthracis pXO1 virulence plasmid encodes a type 1 DNA topoisome  
 A:Reference number: S78064  
 A:Accession: S78064  
 A:Molecule type: DNA  
 A:Residues: 1-784, 'S', 786, 'C', 788-887 <FOW>  
 A:Cross-references: EMBL:M97227; NID:g478995; PIDN:AAA22859.1; PID:g478996  
 A:Experimental source: strain 7702  
 R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle  
 J. Bacteriol. 181, 6509-6515, 1999  
 A>Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor  
 A:Reference number: A59091; MUID:99445483; PMID:10515943  
 A:Accession: E59108  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-784, 'S', 786, 'C', 788-887 <OKI>  
 A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32445.1; PID:g4894357  
 A:Experimental source: strain Sterne  
 A>Note: similar to type 1 DNA topoisomerase, top1, plasmid pXO1; B. anthracis (M97227)  
 C:Genetics:  
 A:Gene: topX; pXO1-142

A:Genome: plasmid  
C:Keywords: DNA binding; isomerase

Query Match 57.4%; Score 39; DB 2; Length 887;  
Best Local Similarity 46.2%; Pred. No. 1e+02;  
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 PIVNLKDEVLSPS 14  
DB 237 PVNKTEDQVLNPS 249  
::: |||::|

RESULT 12  
T04198  
hypothetical protein T4F9.90 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04198  
R:Devan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z15260  
A:Accession: T04198  
A:Molecule type: DNA  
A:Residues: 1-334 <BEV>  
A:Cross-references: UNIPROT:Q9ZSB9; EMBL:AL049523  
A:Experimental source: cultivar Columbia; BAC clone T4F9  
C:Genetics:  
A:Map position: 4  
A:Note: T4F9.90  
C:Superfamily: Arabidopsis thaliana hypothetical protein T4F9.90

Query Match 55.9%; Score 38; DB 2; Length 334;  
Best Local Similarity 56.7%; Pred. No. 47;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PIVNLKDEVLSP 13  
DB 125 PRKLNLDVLPK 136  
| || |||||

RESULT 13  
E88114  
protein F53C3.9 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: E88114  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: AV5000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: E88114  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-485 <STO>  
A:Cross-references: UNIPROT:Q9TXX0; GB:chr\_II; PIDN:AC67460.1; PID:G3786487; GSPDB:GN000  
C:Genetics:  
A:Gene: F53C3.9  
A:Map position: 2  
C:Superfamily: Caenorhabditis elegans ZK1236.4 protein

Query Match 55.9%; Score 38; DB 2; Length 485;  
Best Local Similarity 53.8%; Pred. No. 74;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PIVNLKDEVLSPS 14  
DB 344 PIINYCSEIYSPS 356  
|||::|

RESULT 14  
S68446

origin recognition complex protein, subunit 2 - fission yeast (Schizosaccharomyces pombe  
N:Alternate names: replication factor Orp2  
C:Species: Schizosaccharomyces pombe  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: S68446; T40640  
R:Leatherwood, J.; Lopez-Girona, A.; Russell, P.  
Nature 379, 360-363, 1996  
A:Title: Interaction of Cdc2 and Cdc18 with a fission yeast ORC2-like protein.  
A:Reference number: S68446; MUID:96149252; PMID:8552194  
A:Accession: S68446  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-535 <LEA>  
A:Cross-references: UNIPROT:Q09142; EMBL:U38472; NID:g1185460; PIDN:AAC49165.1; PID:g11  
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, March 1999  
A:Reference number: Z21928  
A:Accession: T40640  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-535 <SEE>  
A:Cross-references: EMBL:AL049474; PIDN:CA839366.1; GSPDB:GNO00067; SPDB:SPBC685.09  
A:Experimental source: strain 972h-; cosmid c685  
C:Genetics:  
A:Gene: ORP2; SPBC685.09  
A:Map position: 2  
C:Function:  
A:Description: interacts with the rate-limiting replication activator Cdc18  
C:Keywords: DNA replication; nucleus

Query Match 55.9%; Score 38; DB 2; Length 535;  
Best Local Similarity 57.1%; Pred. No. 83;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIVNLKDEVLSPS 14  
DB 112 VPLYSLAEPLSPS 125  
|||::|

RESULT 15  
H82516  
60Kda inner-membrane protein XF2780 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82516  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-565 <SIM>  
A:Cross-references: UNIPROT:Q9P9U1; GB:AE004083; GB:AE003849; NID:g9108022; PIDN:AAF855  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, I  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2780  
C:Superfamily: probable 60K inner membrane protein; stage III sporulation protein homol.

Query Match 55.9%; Score 38; DB 2; Length 565;  
Best Local Similarity 54.5%; Pred. No. 88;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPINVKDEVL 11  
||:|:|:|  
Db 74 VPVINVKTDL 84

Search completed: March 10, 2005, 11:10:08  
Job time : 13.1579 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:07:00 ; Search time 85.1053 Seconds  
(without alignments)  
54.260 Million cell updates/sec

Title: US-09-766-412-31

Perfect score: 68

Sequence: 1 VPINLKDEVLSPS 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	14	9	US-09-766-412-31
2	68	100.0	184	14	US-10-131-241-46
3	68	100.0	184	14	US-10-292-418-18
4	68	100.0	184	17	US-10-920-820-13
5	68	100.0	191	9	US-09-998-831-13
6	68	100.0	191	14	US-10-373-561-13
7	68	100.0	207	9	US-09-373-938-2
8	68	100.0	207	13	US-10-080-797-3
9	68	100.0	207	14	US-10-422-934-71
10	59	86.8	43	15	US-10-468-759-2
11	59	86.8	63	9	US-09-822-540A-1
12	59	86.8	63	9	US-09-978-531-5
13	59	86.8	63	9	US-09-978-531-11

14	59	86.8	139	9	US-09-978-531-12	Sequence 12, Appl
15	59	86.8	160	15	US-10-210-172-164	Sequence 164, Appl
16	59	86.8	171	9	US-09-978-531-3	Sequence 3, Appli
17	59	86.8	178	14	US-10-131-241-60	Sequence 60, Appl
18	59	86.8	178	14	US-10-042-347-5	Sequence 5, Appli
19	59	86.8	179	14	US-10-131-241-57	Sequence 57, Appl
20	59	86.8	180	14	US-10-131-241-47	Sequence 47, Appl
21	59	86.8	180	14	US-10-131-241-56	Sequence 56, Appl
22	59	86.8	181	14	US-10-131-241-55	Sequence 55, Appl
23	59	86.8	182	9	US-09-998-831-14	Sequence 14, Appl
24	59	86.8	182	14	US-10-131-241-54	Sequence 54, Appl
25	59	86.8	182	14	US-10-042-347-3	Sequence 3, Appli
26	59	86.8	182	14	US-10-373-561-14	Sequence 14, Appl
27	59	86.8	183	9	US-09-873-676-2	Sequence 2, Appli
28	59	86.8	183	9	US-09-978-531-1	Sequence 1, Appli
29	59	86.8	183	13	US-10-080-797-1	Sequence 1, Appli
30	59	86.8	183	14	US-10-131-241-52	Sequence 52, Appl
31	59	86.8	183	14	US-10-292-418-4	Sequence 4, Appli
32	59	86.8	183	15	US-10-135-872B-11	Sequence 11, Appl
33	59	86.8	183	15	US-10-607-501-2	Sequence 2, Appli
34	59	86.8	183	17	US-10-920-820-12	Sequence 12, Appl
35	59	86.8	184	15	US-10-210-172-162	Sequence 162, Appl
36	59	86.8	184	16	US-10-449-609-4	Sequence 4, Appli
37	59	86.8	185	17	US-10-849-650-4	Sequence 4, Appli
38	59	86.8	208	9	US-09-373-938-5	Sequence 5, Appli
39	59	86.8	220	17	US-10-849-650-2	Sequence 2, Appli
40	59	86.8	385	16	US-10-449-609-6	Sequence 6, Appli
41	59	86.8	682	15	US-10-264-049-3010	Sequence 3010, Ap
42	59	86.8	684	10	US-09-961-403-5	Sequence 5, Appli
43	59	86.8	1516	14	US-10-060-036-166	Sequence 166, Appl
44	59	86.8	1516	15	US-10-431-642-3	Sequence 3, Appli
45	58	85.3	184	10	US-09-938-391-4	Sequence 4, Appli

# ALIGNMENTS

RESULT 1  
US-09-766-412-31  
; Sequence 31, Application US/097666412  
; Patent No. US20020103129A1  
; GENERAL INFORMATION:  
; APPLICANT: GS, Ruowen et al.  
; TITLE OF INVENTION: SMALL PEPTIDES HAVING ANTI-ANGIOGENIC AND ENDOTHELIAL CELL INHIB  
; TITLE OF INVENTION: ACTIVITY  
; FILE REFERENCE: 1781-0215P  
; CURRENT APPLICATION NUMBER: US/09/766,412  
; CURRENT FILING DATE: 2001-01-11  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Endo-3  
US-09-766-412-31

Query Match 100.0%; Score 68; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 VPINLKDEVLSPS 14  
Db 1 VPINLKDEVLSPS 14  
RESULT 2  
US-10-131-241-46  
; Sequence 46, Application US/10131241  
; Publication No. US20030012792A1  
; GENERAL INFORMATION:

; APPLICANT: Holaday, John W.  
; APPLICANT: Fortier, Anne H.  
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation  
; FILE REFERENCE: 05213-0344 43170-271565  
; CURRENT APPLICATION NUMBER: US/10/131,241  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: US 09/413,049  
; PRIOR FILING DATE: 1999-10-06  
; PRIOR APPLICATION NUMBER: US 09/316,802  
; PRIOR FILING DATE: 1999-05-21  
; PRIOR APPLICATION NUMBER: US 60/086,586  
; PRIOR FILING DATE: 1998-05-22  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Murinae sp.  
US-10-131-241-46

Query Match 100.0%; Score 68; DB 14; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
DB 69 VPVNLKDEVLSPS 82

## RESULT 3

US-10-292-418-18  
; Sequence 18, Application US/10292418  
; Publication No. US20030139365A1  
; GENERAL INFORMATION:  
; APPLICANT: Lo, Kin-Ming  
; APPLICANT: Li, Yue  
; APPLICANT: Gillies, Stephen D  
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as  
; FILE REFERENCE: LEX-006C1  
; CURRENT APPLICATION NUMBER: US/10/292,418  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/383,315  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: US 60/097,883  
; PRIOR FILING DATE: 1998-08-25  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-292-418-18

Query Match 100.0%; Score 68; DB 14; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
DB 69 VPVNLKDEVLSPS 82

## RESULT 4

US-10-920-820-13  
; Sequence 13, Application US/10920820  
; Publication No. US20050009144A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer, Inc. and Pfizer Products, Inc.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING DISORDERS  
; FILE REFERENCE: 3153.00234/PC10790A

; CURRENT APPLICATION NUMBER: US/10/920,820  
; CURRENT FILING DATE: 2004-08-17  
; PRIOR APPLICATION NUMBER: US/09/938,391  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US PROV NO. 60/227,924  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: MURINE  
US-10-920-820-13

Query Match 100.0%; Score 68; DB 17; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.00058;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
DB 69 VPVNLKDEVLSPS 82

## RESULT 5

US-09-998-831-13  
; Sequence 13, Application US/09998831  
; Patent No. US20020119153A1  
; GENERAL INFORMATION:  
; APPLICANT: Phillip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; TITLE OF INVENTION: INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998,831  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 09/561,108  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-998-831-13

Query Match 100.0%; Score 68; DB 9; Length 191;  
Best Local Similarity 100.0%; Pred. No. 0.0006;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
DB 76 VPVNLKDEVLSPS 89

## RESULT 6

US-10-373-561-13  
; Sequence 13, Application US/10373561  
; Publication No. US20030175276A1  
; GENERAL INFORMATION:  
; APPLICANT: Phillip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002582  
; CURRENT APPLICATION NUMBER: US/10/373,561  
; CURRENT FILING DATE: 2003-02-24  
; PRIOR APPLICATION NUMBER: US/09/561,499  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0



```
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-10-373-561-13

Query Match      100.0%; Score 68; DB 14; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 76 VPVNLKDEVLSPS 89

RESULT 7
US-09-373-938-2
; Sequence 2, Application US/09373938
; Publication No. US20020115202A1
; GENERAL INFORMATION:
; APPLICANT: Hallenbeck, Paul
; APPLICANT: Chen, Cheayun Theresa
; TITLE OF INVENTION: ADENOVIRAL VECTORS INCLUDING DNA SEQUENCES ENCODING ANGIOGENIC IN
; FILE REFERENCE: 4-308999P1
; CURRENT APPLICATION NUMBER: US/09/373,938
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-373-938-2

Query Match      100.0%; Score 68; DB 9; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 92 VPVNLKDEVLSPS 105

RESULT 8
US-10-080-797-3
; Sequence 3, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Romulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; TITLE OF INVENTION: NEOVASCULARIZATION
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Mouse
US-10-080-797-3

Query Match      100.0%; Score 68; DB 13; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 92 VPVNLKDEVLSPS 105
```

```
RESULT 9
US-10-422-934-71
; Sequence 71, Application US/10422934
; Publication No. US20030186841A1
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F., III
; APPLICANT: Kadan, Michael
; APPLICANT: Beerli, Roger
; TITLE OF INVENTION: LIGAND ACTIVATED TRANSCRIPTIONAL REGULATOR PROTEINS
; FILE REFERENCE: 22908-1227C
; CURRENT APPLICATION NUMBER: US/10/422,934
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/586,625
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 09/433,042
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Muridae
US-10-422-934-71

Query Match      100.0%; Score 68; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 92 VPVNLKDEVLSPS 105
```

```
RESULT 10
US-10-468-759-2
; Sequence 2, Application US/10468759
; Publication No. US20040073007A1
; GENERAL INFORMATION:
; APPLICANT: Universita degli Studi di Milano
; TITLE OF INVENTION: Peptides with angiogenic activity
; FILE REFERENCE: 683pct
; CURRENT APPLICATION NUMBER: US/10/468,759
; CURRENT FILING DATE: 2003-08-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-468-759-2

Query Match      86.8%; Score 59; DB 15; Length 43;
Best Local Similarity 85.7%; Pred. No. 0.0044;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
Db 20 VPVNLKDEVLSPS 33
```

```
RESULT 11
US-09-822-540A-1
; Sequence 1, Application US/09822540A
; Patent No. US2002058620A1
; GENERAL INFORMATION:
; APPLICANT: Phoenix Pharmaceuticals, Inc.
; APPLICANT: Chang, Jaw-Kang
; TITLE OF INVENTION: Cell Growth Regulation System
; FILE REFERENCE: PhoenixPharFull
; CURRENT APPLICATION NUMBER: US/09/822,540A
; CURRENT FILING DATE: 2001-03-30
```

; PRIOR APPLICATION NUMBER: 60/194,561  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-822-540A-1

Query Match 86.8%; Score 59; DB 9; Length 63;  
Best Local Similarity 85.7%; Pred. No. 0.0068;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
| | | | | | | | | | | | | | | | | |  
Db 6 VPVNLKDELLFPS 19

## RESULT 12

US-09-978-531-5  
; Sequence 5, Application US/09978531  
; Publication No. US20020086007A1  
; GENERAL INFORMATION:

; APPLICANT: Sim, Kim Lee  
; APPLICANT: Liang, Hong  
; TITLE OF INVENTION: Angiogenesis-Inhibiting Peptides and Proteins and Methods of Use  
; FILE REFERENCE: 05213-0811 43170-264992  
; CURRENT APPLICATION NUMBER: US/09/978,531  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/240,127  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-531-5

Query Match 86.8%; Score 59; DB 9; Length 63;  
Best Local Similarity 85.7%; Pred. No. 0.0068;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
| | | | | | | | | | | | | | | | | |  
Db 6 VPVNLKDELLFPS 19

## RESULT 13

US-09-978-531-11  
; Sequence 11, Application US/09978531  
; Publication No. US20020086007A1  
; GENERAL INFORMATION:

; APPLICANT: Sim, Kim Lee  
; APPLICANT: Liang, Hong  
; TITLE OF INVENTION: Angiogenesis-Inhibiting Peptides and Proteins and Methods of Use  
; FILE REFERENCE: 05213-0811 43170-264992  
; CURRENT APPLICATION NUMBER: US/09/978,531  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/240,127  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 63  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-531-11

Query Match 86.8%; Score 59; DB 9; Length 63;  
Best Local Similarity 85.7%; Pred. No. 0.0068;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
| | | | | | | | | | | | | | | | | |  
Db 6 VPVNLKDELLFPS 19

## RESULT 14

US-09-978-531-12  
; Sequence 12, Application US/09978531  
; Publication No. US20020086007A1  
; GENERAL INFORMATION:

; APPLICANT: Sim, Kim Lee  
; APPLICANT: Liang, Hong  
; TITLE OF INVENTION: Angiogenesis-Inhibiting Peptides and Proteins and Methods of Use  
; FILE REFERENCE: 05213-0811 43170-264992  
; CURRENT APPLICATION NUMBER: US/09/978,531  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: US 60/240,127  
; PRIOR FILING DATE: 2000-10-13  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-531-12

Query Match 86.8%; Score 59; DB 9; Length 139;  
Best Local Similarity 85.7%; Pred. No. 0.017;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14  
| | | | | | | | | | | | | | | | | |  
Db 25 VPVNLKDELLFPS 38

## RESULT 15

US-10-210-172-164  
; Sequence 164, Application US/10210172  
; Publication No. US20040043928A1  
; GENERAL INFORMATION:

; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Miller, Charles  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Pena, Carol  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Shinkete, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Li, Li  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Casman, Stacie  
; APPLICANT: Voss, Edward  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Gorman, Linda  
; APPLICANT: Leite, Mario  
; APPLICANT: Vernet, Corine  
; APPLICANT: Anderson, David  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Hjal, Tord  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Alsobrook II, John  
; APPLICANT: Lepley, Denise et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-416 A

Fri Mar 11 11:21:03 2005

```

; CURRENT APPLICATION NUMBER: US/10/210,172
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/323,994
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/373,814
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/310,544
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/312,892
; PRIOR FILING DATE: 2001-08-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 327
; SOFTWARE: CurSeqList version 0.1
; SEQ ID NO 164
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-172-164

Query Match      86.8%; Score 59; DB 15; Length 160;
Best Local Similarity 85.7%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 VPVNLKDEVLSPS 14
Db      44 VPVNLKDELLFPS 57
```

Search completed: March 10, 2005, 11:15:27  
Job time : 86.1053 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:04:09 ; Search time 28.3684 Seconds  
(without alignments)  
36.840 Million cell updates/sec

Title: US-09-766-412-31

Perfect score: 68  
Sequence: 1 VPVNLKDEVLSPS 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	14	3	US-09-385-442-31
2	68	100.0	184	4	US-09-938-391A-13
3	68	100.0	191	3	US-09-561-500-13
4	68	100.0	191	3	US-09-561-108-13
5	68	100.0	191	4	US-09-561-526-13
6	68	100.0	191	4	US-09-561-499-13
7	68	100.0	191	4	US-09-998-831-13
8	68	100.0	191	4	US-09-561-005-13
9	68	100.0	195	1	US-08-159-784-2
10	59	86.8	63	4	US-09-822-540A-1
11	59	86.8	178	3	US-09-315-689-5
12	59	86.8	182	3	US-09-561-500-14
13	59	86.8	182	3	US-09-561-108-14
14	59	86.8	182	3	US-09-315-689-3
15	59	86.8	182	4	US-09-561-526-14
16	59	86.8	182	4	US-09-561-499-14
17	59	86.8	182	4	US-09-998-831-14
18	59	86.8	182	4	US-09-561-005-14
19	59	86.8	182	3	US-09-206-059-2
20	59	86.8	183	4	US-09-938-391A-12
21	59	86.8	684	4	US-09-961-403-5
22	59	86.8	1516	4	US-09-949-016-8209
23	58	85.3	184	4	US-09-938-391A-2
24	58	85.3	230	4	US-09-938-391A-2
25	54.5	80.1	185	3	US-08-985-526-36
26	50	73.5	184	4	US-09-938-391A-11
27	41	60.3	222	4	US-09-602-787A-298

28	41	60.3	222	4	US-09-602-787A-300	Sequence 300, App
29	40	58.8	181	1	US-09-589-483-20	Sequence 20, Appl
30	40	58.8	191	1	US-08-159-784-3	Sequence 3, Appli
31	40	58.8	548	3	US-08-656-034-10	Sequence 10, Appl
32	40	58.8	628	3	US-08-656-034-2	Sequence 2, Appli
33	39	57.4	750	4	US-09-489-039A-8940	Sequence 8940, Ap
34	38	55.9	99	4	US-09-134-000C-5491	Sequence 5491, Ap
35	38	55.9	771	2	US-08-742-753-2	Sequence 2, Appli
36	37	54.4	693	4	US-09-248-796A-18167	Sequence 18167, A
37	36	52.9	100	4	US-09-583-110-3743	Sequence 3743, Ap
38	36	52.9	186	4	US-09-583-110-3896	Sequence 3896, Ap
39	36	52.9	190	4	US-09-107-433-3585	Sequence 3585, Ap
40	36	52.9	365	4	US-09-902-540-16239	Sequence 16239, A
41	36	52.9	427	4	US-09-949-016-8279	Sequence 8279, Ap
42	36	52.9	505	4	US-09-328-352-4724	Sequence 4724, Ap
43	36	52.9	509	4	US-09-902-540-11469	Sequence 11469, A
44	36	52.9	755	4	US-09-248-796A-15109	Sequence 15109, A
45	35	51.5	57	4	US-09-270-767-32217	Sequence 32217, A

## ALIGNMENTS

### RESULT 1

US-09-385-442-31  
; Sequence 31, Application US/09385442  
; Patent No. 6200954  
; GENERAL INFORMATION:  
; APPLICANT: Ge, Ruowen  
; APPLICANT: Kini, R. Manjunatha  
; TITLE OF INVENTION: Small Peptides Having Potent Anti-Angiogenic Activity  
; FILE REFERENCE: 1781-170P  
; CURRENT APPLICATION NUMBER: US/09/385,442  
; CURRENT FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 60/099,313  
; EARLIER FILING DATE: 1999-09-04  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 31  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: mammalian  
; FEATURE:  
; OTHER INFORMATION: Endo-3  
US-09-385-442-31

Query Match 100.0%; Score 68; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 6.8e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

### QY

1 VPVNLKDEVLSPS 14

### Db

1 VPVNLKDEVLSPS 14

### RESULT 2

US-09-938-391A-13  
; Sequence 13, Application US/09938391A  
; Patent No. 6803211  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer, Inc. and Pfizer Products, Inc..  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING DISORDERS INVOLVING ANGIOGENESIS  
; FILE REFERENCE: 3153.00234/PC10790A  
; CURRENT APPLICATION NUMBER: US/09/938,391A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US PROV NO. 60/227,924  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 184  
; TYPE: PRT

; ORGANISM: MURINE  
 US-09-938-391A-13  
  
 Query Match 100.0%; Score 68; DB 4; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 VPVNLKDEVLSPS 14  
 Db 69 VPVNLKDEVLSPS 82  
 |||||  
  
 RESULT 3  
 US-09-561-500-13  
 ; Sequence 13, Application US/09561500  
 ; Patent No. 6342219  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Philip E. Thorpe  
 ; APPLICANT: Rolf A. Brekken  
 ; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
 ; FILE REFERENCE: 4001.002500  
 ; CURRENT APPLICATION NUMBER: US/09/561,500  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/131,432  
 ; PRIOR FILING DATE: 1999-04-28  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 191  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
 US-09-561-500-13  
  
 Query Match 100.0%; Score 68; DB 3; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 VPVNLKDEVLSPS 14  
 Db 76 VPVNLKDEVLSPS 89  
 |||||  
  
 RESULT 4  
 US-09-561-108-13  
 ; Sequence 13, Application US/09561108  
 ; Patent No. 6342221  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Philip E. Thorpe  
 ; APPLICANT: Rolf A. Brekken  
 ; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
 ; FILE REFERENCE: 4001.002584  
 ; CURRENT APPLICATION NUMBER: US/09/561,108  
 ; CURRENT FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/131,432  
 ; PRIOR FILING DATE: 1999-04-28  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 191  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
 US-09-561-108-13  
  
 Query Match 100.0%; Score 68; DB 3; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 VPVNLKDEVLSPS 14  
 Db 76 VPVNLKDEVLSPS 89  
 |||||  
  
 RESULT 5  
 US-09-561-526-13  
 ; Sequence 13, Application US/09561526  
 ; Patent No. 6416758  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Philip E. Thorpe  
 ; APPLICANT: Rolf A. Brekken  
 ; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF  
 ; FILE REFERENCE: 4001.002586  
 ; CURRENT APPLICATION NUMBER: US/09/561,526  
 ; CURRENT FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/131,432  
 ; PRIOR FILING DATE: 1999-04-28  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 191  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
 US-09-561-526-13  
  
 Query Match 100.0%; Score 68; DB 4; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 VPVNLKDEVLSPS 14  
 Db 76 VPVNLKDEVLSPS 89  
 |||||  
  
 RESULT 6  
 US-09-561-499-13  
 ; Sequence 13, Application US/09561499  
 ; Patent No. 6524583  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Philip E. Thorpe  
 ; APPLICANT: Rolf A. Brekken  
 ; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
 ; FILE REFERENCE: 4001.002582  
 ; CURRENT APPLICATION NUMBER: US/09/561,499  
 ; CURRENT FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: 60/131,432  
 ; PRIOR FILING DATE: 1999-04-28  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 191  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
 US-09-561-499-13  
  
 Query Match 100.0%; Score 68; DB 4; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 0.00013;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 VPVNLKDEVLSPS 14  
 Db 76 VPVNLKDEVLSPS 89  
 |||||  
  
 RESULT 7  
 US-09-998-831-13  
 ; Sequence 13, Application US/09998831  
 ; Patent No. 6676941  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Philip E. Thorpe

```
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
; FILE OF INVENTION: INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/998,831
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/561,108
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-998-831-13

Query Match 100.0%; Score 68; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
DB 76 VPVNLKDEVLSPS 89

RESULT 8
US-09-561-005-13
; Sequence 13, Application US/09561005
; Patent No. 6703020
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; TITLE OF INVENTION: ANTIBODY CONJUGATE METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002585
; CURRENT APPLICATION NUMBER: US/09/561,005
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-005-13

Query Match 100.0%; Score 68; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
DB 76 VPVNLKDEVLSPS 89

RESULT 9
US-08-159-784-2
; Sequence 2, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
```

```
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-159-784-2

Query Match 100.0%; Score 68; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
DB 80 VPVNLKDEVLSPS 93

RESULT 10
US-09-822-540A-1
; Sequence 1, Application US/09822540A
; Patent No. 6835806
; GENERAL INFORMATION:
; APPLICANT: Phoenix Pharmaceuticals, Inc.
; APPLICANT: Chang, Jaw-Kang
; TITLE OF INVENTION: Cell Growth Regulation System
; FILE REFERENCE: PhoenixPharFull
; CURRENT APPLICATION NUMBER: US/09/822,540A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/194,561
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-540A-1

Query Match 86.8%; Score 59; DB 4; Length 63;
Best Local Similarity 85.7%; Pred. No. 0.0016;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPVNLKDEVLSPS 14
DB 6 VPVNLKDEVLSPS 19

RESULT 11
US-09-315-689-5
; Sequence 5, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
```

```

; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-5

Query Match      86.8%; Score 59; DB 3; Length 178;
Best Local Similarity 85.7%; Pred. No. 0.0052;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPIVNLKDEVLSPS 14
   |||||:|:|:|
Db 65 VPIVNLKDELLFPS 78

RESULT 12
US-09-561-500-14
; Sequence 14, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-561-500-14

Query Match      86.8%; Score 59; DB 3; Length 182;
Best Local Similarity 85.7%; Pred. No. 0.0053;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPIVNLKDEVLSPS 14
   |||||:|:|:|
Db 69 VPIVNLKDELLFPS 82

RESULT 13
US-09-561-108-14
; Sequence 14, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-561-108-14

Query Match      86.8%; Score 59; DB 3; Length 182;
Best Local Similarity 85.7%; Pred. No. 0.0053;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPIVNLKDEVLSPS 14
   |||||:|:|:|
Db 69 VPIVNLKDELLFPS 82

RESULT 14
US-09-315-689-3
; Sequence 3, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-3

Query Match      86.8%; Score 59; DB 3; Length 182;
Best Local Similarity 85.7%; Pred. No. 0.0053;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPIVNLKDEVLSPS 14
   |||||:|:~|:|
Db 69 VPIVNLKDELLFPS 82

RESULT 15
US-09-561-526-14
; Sequence 14, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-561-526-14

Query Match      86.8%; Score 59; DB 4; Length 182;
Best Local Similarity 85.7%; Pred. No. 0.0053;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPIVNLKDEVLSPS 14
   |||||:|:~|:|
Db 69 VPIVNLKDELLFPS 82

```



Db           |||||||:|  
             69 VPIVNLKDELLPES 82

Search completed: March 10, 2005, 11:11:31  
Job time : 29.3684 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 10:55:58 ; Search time 71.1053 Seconds  
(without alignments)  
76.150 Million cell updates/sec

Title: US-09-766-412-31

Perfect score: 68  
Sequence: 1 VFIVNLKDEVLSFS 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	68	100.0	14	4	AAB74257	Aab74257 Exemplary
2	68	100.0	14	4	AAB80863	Aab80863 Angiostat
3	68	100.0	14	5	ABG97540	Abg97540 Antiangio
4	68	100.0	50	4	AAB35585	Aab35585 Antiangio
5	68	100.0	108	4	AAB49807	Aab49807 Murine en
6	68	100.0	183	5	AAM49504	Aam49504 Mouse end
7	68	100.0	184	2	AAV18409	Aay18409 Endostati
8	68	100.0	184	2	AAV08689	Aay08689 Murine en
9	68	100.0	184	3	AAV70258	Aay70258 Murine an
10	68	100.0	184	4	AAB49380	Aab49380 Murine en
11	68	100.0	184	5	ABG31793	Abg31793 Human end
12	68	100.0	184	7	ABU64290	Abu64290 Murine en
13	68	100.0	191	3	AAB28398	Aab28398 Murine en
14	68	100.0	191	5	AAY77950	Aau77950 Amino aci
15	68	100.0	207	4	AAE02031	Aae02031 Murine en
16	68	100.0	207	4	AAB71930	Aab71930 Murine en
17	68	100.0	207	5	ABY79902	Abb79902 Mouse end
18	68	100.0	207	8	ADN00602	Adn00602 Murine en
19	68	100.0	207	8	ADO43904	Ado43904 Amino aci
20	68	100.0	218	2	AAV08691	Aay08691 Murine ge
21	68	100.0	580	2	AAV08692	Aay08692 Murine ge
22	68	100.0	684	2	AAV25114	Aay25114 Mouse alp
23	68	100.0	1288	2	AAW26328	Aaw26328 Mouse alp
24	68	100.0	1288	2	AAW92297	Aaw92297 Mouse alp
25	68	100.0	1315	8	ABO84585	Abo84585 Mouse can

26	68	100.0	1527	8	ABO84584	Abo84584 Mouse can
27	59	86.8	43	5	ABG31788	Abg31788 Human end
28	59	86.8	63	5	ABR08121	Abbr08121 Human end
29	59	86.8	160	7	ADe40258	Ade40258 Human NOV
30	59	86.8	178	3	AAy94324	Aay94324 Alternate
31	59	86.8	178	4	AAU00900	Aau00900 Human end
32	59	86.8	179	4	AAU00901	Aau00901 Human end
33	59	86.8	180	4	AAU00899	Aau00899 Human end
34	59	86.8	181	4	AAU00898	Aau00898 Human end
35	59	86.8	182	3	AAV59622	Aay59622 Human end
36	59	86.8	182	3	AAy94323	Aay94323 Human end
37	59	86.8	182	3	AAB28399	Aab28399 Human end
38	59	86.8	182	4	AAU00897	Aau00897 Human end
39	59	86.8	182	5	AAU77951	Aau77951 Amino aci
40	59	86.8	183	2	AAV02113	Aay02113 SPQ ID 76
41	59	86.8	183	2	AAV08693	Aay08693 Human end
42	59	86.8	183	3	AAV70252	Aay70252 Human ang
43	59	86.8	183	3	AAI90771	Aay90771 Human ang
44	59	86.8	183	3	AAB16451	Aab16451 Human end
45	59	86.8	183	3	AAB30493	Aab30493 Amino aci

ALIGNMENTS

RESULT 1  
AAB74257  
ID AAB74257 standard; peptide; 14 AA.  
XX AC AAB74257;  
XX DT 23-MAY-2001 (first entry)  
XX DE Exemplary anti-angiogenic peptide #31.  
XX KW Anti-angiogenesis; plasminogen; VEGF; vascular endothelial growth factor;  
XX LW FLT-1; tumour; metastasis; cancer.  
XX OS Homo sapiens.  
XX PN WO200118030-A2.  
XX PD 15-MAR-2001.  
XX PF 01-SEP-2000; 2000WO-SG000131.  
XX PR 03-SEP-1999; 99SG-00004310.  
XX (UYSI-) UNIV SINGAPORE NAT.  
XX Ge R, Kini RM;  
XX WPI; 2001-257785/26.  
XX Peptides comprising a portion of a protein selected from plasminogen, endostatin, VEGF, FLT-1 and KDR/FLK-1 are useful for treating primary tumor growth.  
XX Claim 7; Page 18; 34pp; English.  
XX The present invention relates to anti-angiogenesis peptides from a portion of a selected from plasminogen, endostatin, VEGF or FLT-1. The invention is used to prevent or treat primary tumour growth or metastasis or undesired angiogenesis  
XX Sequence 14 AA;  
Query Match 100.0%; Score 68; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY : 1 VFIVNLKDEVLSFS 14  
|||||

Db. 1 VPVNLKDEVLSPS 14

RESULT 2  
ID AAB80863 standard; peptide; 14 AA.  
XX AAB80863;  
AC AAB80863;  
DT 29-MAY-2001 (first entry)  
XX Angiostatin-derived peptide #20.  
DE Cytostatic; angiogenesis inhibitor; tumour metastasis; tumour growth.  
KW Unidentified.  
XX OS  
XX US6200954-B1.  
PN 13-MAR-2001.  
XX 30-AUG-1999; 99US-00385442.  
XX 04-SEP-1998; 98US-0099313P.  
PR (UYSI-) UNIV SINGAPORE NAT.  
PA Ge R, Kini RM;  
PI WPI; 2001-234520/24.  
XX Peptides having potent anti-angiogenic activity, useful for the treatment  
PT of tumors, comprise a portion of a Flt-1 protein.  
XX Disclosure; Col 15-16; 21pp; English.  
XX The present sequence is a peptide which is effective in inhibiting  
CC undesirable angiogenesis. Angiogenesis is the process of new blood vessel  
CC formation from pre-existing vessels. Inappropriate angiogenesis is  
CC associated with various pathological conditions including solid tumour  
CC growth and metastasis. The present peptide can be used to prevent  
CC tumour metastasis or inhibit the growth of a primary tumour  
XX  
SQ Sequence 14 AA;  
Query Match 100.0%; Score 68; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VPVNLKDEVLSPS 14  
DB 1 VPVNLKDEVLSPS 14

RESULT 3  
AAB80863  
ID AAB80863 standard; peptide; 14 AA.  
XX AAB80863;  
AC AAB80863;  
DT 16-DEC-2002 (first entry)  
XX Antiangiogenic peptide, Endo-3.  
DE Angiogenesis; plasminogen; endostatin; Endo-; VEGF;  
KW vascular endothelial growth factor; Flt-1; kinase insert domain; Flt-1;  
KW angiogenic inhibitor; Angio-; endothelial cell; proliferation;  
KW tumour growth; blood vessel formation; migration; tubule structure;  
KW embryonic development; wound healing; tumour metastasis;  
KW rheumatoid arthritis; psoriasis; anticancer; therapy;  
XX antiangiogenic therapy; mitogen; tyrosine kinase receptor.  
XX Mammalia.

XX US2002103129-A1.  
XX 01-AUG-2002.  
XX 22-JAN-2001; 2001US-00766412.  
XX 04-SEP-1998; 98US-0099313P.  
XX 30-AUG-1999; 99US-00385442.  
XX (GERR/) GE R.  
XX (KINI/) KINI R M.  
XX Ge R, Kini RM;  
XX WPI; 2002-749595/81.  
XX Novel peptide for preventing/treating undesired angiogenesis, has a  
XX portion of a protein e.g. plasminogen, endostatin, and potent  
XX antiangiogenic activity and endothelial cell proliferation inhibition  
XX activity.  
XX Claim 9; Page 10; 24pp; English.  
XX The invention discloses a peptide comprising a portion of a protein e.g.  
XX plasminogen, endostatin (Endo-1-4), vascular endothelial growth factor  
XX (VEGF), Flt-1 or kinase insert domain containing receptor/Flt-1. These  
XX peptides are able to act as potent angiogenic inhibitors (Angio-1-5)  
XX inhibiting endothelial cell proliferation and retarding tumour growth.  
XX Angiogenesis is the process of new blood vessel formation from pre-  
XX existing vessel, involving endothelial cell proliferation and wound healing.  
XX physiological functions such as embryonic development and wound healing.  
XX Inappropriate angiogenesis is also associated with various pathological  
XX conditions including tumour growth and metastasis, rheumatoid arthritis  
XX and psoriasis. Anticancer therapy by inhibiting tumour angiogenesis is  
XX called antiangiogenic therapy. VEGF is an endothelial specific mitogen  
XX that functions through two high affinity tyrosine kinase receptors, Flt-1  
XX and Flk-1 (not defined). Protein-protein interactions are crucial to many  
XX physiological and pharmacological processes and the peptides disclosed  
XX interfere with these interactions. The peptides are useful for preventing  
XX or treating undesired angiogenesis and primary tumour growth or  
XX metastasis. The sequences presented in ABG97510-ABG97559 are examples of  
XX antiangiogenic peptides  
XX  
SQ Sequence 14 AA;  
Query Match 100.0%; Score 68; DB 5; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.6e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VPVNLKDEVLSPS 14  
DB 1 VPVNLKDEVLSPS 14

RESULT 4  
AAB35585  
ID AAB35585 standard; peptide; 50 AA.  
XX AAB35585;  
AC AAB35585;  
DT 14-FEB-2001 (first entry)  
XX Antiangiogenic pentapeptide II.  
DE Antiangiogenic; angiogenesis; cancer; endostatin.  
KW Synthetic.  
XX WO200063249-A1.  
XX 26-OCT-2000.

XX 11-APR-2000; 2000WO-EP003236.  
 XX 15-APR-1999; 99IT-MI000777.  
 XX (UYMI-) UNIV MILANO.  
 XX (UYFI-) UNIV FIRENZE.  
 XX Chillemi F, Francescato P, Ziche M;  
 XX WPI; 2001-007005/01.  
 XX Polypeptides derived from endostatin exhibiting antiangiogenic activity  
 XX useful for treatment of angiogenesis-dependent tumors.  
 XX Claim 3; Page 18; 28pp; English.  
 XX The present invention describes a number of peptides derived from  
 XX endostatin which exhibit antiangiogenic activity. These may be used in  
 XX the treatment of cancer. The present sequence is one of the peptides of  
 XX the invention  
 XX Sequence 50 AA;  
 XX Query Match 100.0%; Score 68; DB 4; Length 50;  
 XX Best Local Similarity 100.0%; Pred. No. 0.00016;  
 XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 DB 30 VPVNLKDEVLSPS 43  
 RESULT 5  
 AAB49807  
 ID AAB49807 standard; protein; 108 AA.  
 XX AC AAB49807;  
 XX DT 02-MAR-2001 (first entry)  
 XX DE Murine endostatin peptide fragment SEQ ID NO: 20.  
 XX Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken; cancer;  
 XX inflammation; angiogenesis-dependent disease.  
 XX Mus musculus.  
 XX WO200067771-A1.  
 XX PD 16-NOV-2000.  
 XX PF 02-MAY-2000; 2000WO-US012063.  
 XX PR 06-MAY-1999; 99US-0132907P.  
 XX PR 14-JUL-1999; 99US-00353333.  
 XX (BURN-) BURNHAM INST.  
 XX Vuori K;  
 XX WPI; 2001-040937/05.  
 XX Endostatin peptide comprising at least four endostatin amino acid  
 XX residues are e.g. angiogenesis inhibitors for treating cancer and  
 XX diabetic retinopathy.  
 XX Claim 10; Page 124; 146pp; English.  
 XX The present invention provides endostatin peptides which can be used in  
 XX the modulation of angiogenesis. This is useful in the treatment of  
 XX cancers, inflammation, rheumatoid arthritis, chronic articular  
 XX rheumatism, psoriasis, disorders associated with inopportune invasion of

CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
 CC of prematurity, macular degeneration, corneal graft rejection,  
 CC retrolental fibroplasia, rubeosis, capillary proliferation in  
 CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
 CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints and wound  
 CC granulation. In addition, the peptides can be used as birth control  
 CC agents  
 XX Sequence 108 AA;  
 SQ Query Match 100.0%; Score 68; DB 4; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 0.00038;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 DB 46 VPVNLKDEVLSPS 59  
 RESULT 6  
 AAM49504  
 ID AAM49504 standard; protein; 183 AA.  
 XX AC AAM49504;  
 XX DT 07-MAY-2002 (first entry)  
 XX DE Mouse endostatin protein.  
 XX Endostatin; murine; proliferation; blood vessel endothelium;  
 XX regeneration; tumour; blood vessel; treatment; amplification.  
 XX Mus sp.  
 XX Key Location/Qualifiers  
 XX FT Misc-difference 181  
 XX FT /label= Ser, Phe  
 XX CN1177005-A.  
 XX PD 25-MAR-1998.  
 XX PF 10-SEP-1997; 97CN-00107112.  
 XX PR 10-SEP-1997; 97CN-00107112.  
 XX (XUGG/) XU G.  
 XX Xu G, Ren M, Xu L;  
 XX WPI; 2002-106746/15.  
 XX Gene clone of inhibitory factor for hyperplasia of inner blood vessel  
 XX cells in human body's real tumor, and its use in anti-tumor blood vessel  
 XX regeneration.  
 XX Disclosure; Page 4 (Disclosure); 6pp; Chinese.  
 XX This invention describes a novel preparation which inhibits the  
 XX proliferation of blood vessel endothelium and prevents the regeneration  
 XX activity of tumour blood vessels. The preparation can also be used as a  
 XX biological preparation in the treatment of tumours. This sequence  
 XX represents the murine endostatin protein described in the invention  
 XX Sequence 183 AA;  
 SQ Query Match 100.0%; Score 68; DB 5; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 0.00069;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 DB 46 VPVNLKDEVLSPS 59

Murine; immunoglobulin Fc fragment; endostatin; immunofusin;  
KW angiogenesis; inhibitor; cytostatic; anti-rheumatoid; anti-arthritis;  
KW antipsoriatic; antidiabetic; ophthalmological; immunosuppressant;  
KW vasotonic; vulnerable; treatment; antiarteriosclerosis; tumour;

KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;  
 KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;  
 KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;  
 KW wound granulation; keloid scar; gene therapy.  
 XX Mus musculus.  
 XX WO200011033-A2.  
 XX PD 02-MAR-2000.  
 XX PF 25-AUG-1999; 99WO-US019329.  
 XX PR 25-AUG-1998; 98US-0097883P.  
 XX PA (LEXI-) LEXINGEN PHARM CORP.  
 XX PI Lo K, Li Y, Gillies SD;  
 XX WI; 2000-237616/20.  
 XX DR N-PSDB; AA251299.  
 XX PT Novel fusion protein of angiostatin or endostatin and an immunoglobulin  
 PT FC region, useful for treating conditions mediated by angiogenesis, such  
 PT as rheumatoid arthritis, tumors and macular degeneration.  
 XX Example 5; Page 48-49; 68pp; English.  
 XX CC The patent discloses a DNA molecule encoding a fusion protein comprising  
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis  
 CC inhibitor selected from angiostatin, endostatin, a plasminogen fragment  
 CC having angiostatin activity, a collagen XVIII fragment having endostatin  
 CC activity, or combinations of them. The fusion protein (immunofusin) is  
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated  
 CC by angiogenesis. Conditions that may be treated include solid tumours,  
 CC blood born tumours, tumour metastasis, benign tumours including  
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic  
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases  
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular  
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental  
 CC fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,  
 CC plaque neovascularisation, telangiectasia, haemophilic joints,  
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of  
 CC endothelial cells, intestinal cells, atherosclerosis, sclerodermal and  
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used in  
 CC gene therapy. The present sequence is a murine endostatin used in the  
 CC construction of immunofusin containing murine immunoglobulin Fc fragment  
 XX SQ Sequence 184 AA;  
 Query Match 100.0%; Score 68; DB 3; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 DB 69 VPVNLKDEVLSPS 82  
 RESULT 10  
 ID AAB49380  
 AC AAB49380 standard; protein; 184 AA.  
 XX AAB49380;  
 XX 02-MAR-2001 (first entry)  
 XX Murine endostatin SEQ ID NO: 4.  
 XX Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken; cancer;  
 KW inflammation; angiogenesis-dependent disease.  
 XX Mus musculus.

XX WO200067771-A1.  
 XX PD 16-NOV-2000.  
 XX PF 02-MAY-2000; 2000WO-US012063.  
 XX PR 06-MAY-1999; 99US-0132907P.  
 XX PR 14-JUL-1999; 99US-00353333.  
 XX PA (BURN-) BURNHAM INST.  
 XX PI Vuori K;  
 XX WI; 2001-040937/05.  
 XX DR N-PSDB; AAC88290.  
 XX PT Endostatin peptide comprising at least four endostatin amino acid  
 PT residues are e.g. angiogenesis inhibitors for treating cancer and  
 PT diabetic retinopathy.  
 XX PS Disclosure; Fig 1; 146pp; English.  
 XX CC The present invention provides endostatin peptides which can be used in  
 CC the modulation of angiogenesis. This is useful in the treatment of  
 CC cancers, inflammation, rheumatoid arthritis, chronic articular  
 CC rheumatism, psoriasis, disorders associated with inopportune invasion of  
 CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
 CC of prematurity, macular degeneration, corneal graft rejection,  
 CC retrolental fibroplasia, rubeosis, capillary proliferation in  
 CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
 CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints and wound  
 CC granulation. In addition, the peptides can be used as birth control  
 CC agents  
 XX SQ Sequence 184 AA;  
 Query Match 100.0%; Score 68; DB 4; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 DB 69 VPVNLKDEVLSPS 82  
 RESULT 11  
 ID ABG31793  
 AC ABG31793 standard; protein; 184 AA.  
 XX ABG31793;  
 XX 05-DEC-2002 (first entry)  
 XX Human endostatin polypeptide.  
 XX Human; endostatin; tumour; cancer; metastasis; cytostatic;  
 KW antiangiogenic.  
 XX Homo sapiens.  
 XX WO200268457-A2.  
 XX 06-SEP-2002.  
 XX 27-FEB-2002; 2002WO-IT000119.  
 XX 27-FEB-2001; 2001IT-MI000394.  
 XX (UYMI-) UNIV MILANO.  
 XX Chillemi F, Vicentinie LMT, Francescato P;

XX WPI; 2002-698655/75.  
 XX New peptide useful for the preparation of medicaments with antiangiogenic  
 PT activity that may be used in treating tumors or metastases, comprises a  
 PT sequence corresponding to fragments of human endostatin.  
 XX  
 PS Disclosure; Fig 1; 24pp; English.  
 XX  
 CC The invention relates to peptide comprising 20-50 amino acids with  
 CC sequences corresponding to the human endostatin polypeptide sequence, its  
 CC salt or non-toxic derivative. The peptides are useful in the preparation  
 CC of medicaments with antiangiogenic activity which may be useful in  
 CC treating tumors or metastases. This sequence represents a human  
 CC endostatin polypeptide  
 XX  
 SQ Sequence 184 AA;  
 Query Match 100.0%; Score 68; DB 5; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 69 VPVNLKDEVLSPS 82  
 |||||  
 RESULT 12  
 ABU64290  
 ID ABU64290 standard; protein; 184 AA.  
 XX  
 AC ABU64290;  
 XX  
 DT 11-MAR-2004 (first entry)  
 XX  
 DE Murine endostatin protein.  
 XX  
 KW Vector; rAAV; recombinant adeno-associated viral vector;  
 KW anti-angiogenesis; PDF; angiogenesis; eye disorder; blindness;  
 KW retinal degeneration; macular degeneration; neovascularisation;  
 KW ophthalmological.  
 XX  
 OS Mus sp.  
 XX  
 FN WO2003080648-A2.  
 XX  
 PD 02-OCT-2003.  
 XX  
 PF 20-MAR-2003; 2003WO-US008667.  
 XX  
 PR 20-MAR-2002; 2002US-0366114P.  
 XX  
 PA (UYFL) UNIV FLORIDA RES FOUND INC.  
 PA (UYJO) UNIV JOHNS HOPKINS.  
 XX  
 PI Hauswirth WW, Campochiaro PA, Berns KI;  
 XX  
 DR WPI; 2003-779243/73.  
 DR N-PSDB; AAL56259.  
 XX  
 PT Novel adeno-associated viral vector comprising polynucleotide encoding  
 PT pigment epithelium-derived factor, useful for treating choroidal  
 PT neovascularization, blindness, loss of vision.  
 XX  
 PS Claim 14; Page 38; Opp; English.  
 XX  
 CC The present invention relates to an adeno-associated viral (AAV) vector  
 CC comprising a polynucleotide that comprises a nucleic acid segment that  
 CC encodes a choroidal or ocular neovascularisation inhibitory polypeptide  
 CC operably linked to a promoter that expresses the segment to produce the  
 CC polypeptide in a selected mammalian host cell. Such a vector is useful  
 CC for providing a choroidal or ocular neovascularisation inhibitory  
 CC polypeptide to a mammal, for use in the therapy of ocular

CC neovascularisation, choroidal neovascularisation, retinal  
 CC neovascularisation, age-related macular degeneration, visual impairment,  
 CC ocular dysfunction, loss of vision, retinopathy, or blindness in a human.  
 CC The present sequence is a protein shown in the exemplification of the  
 CC invention  
 XX  
 SQ Sequence 184 AA;  
 Query Match 100.0%; Score 68; DB 7; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 0.0007;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 69 VPVNLKDEVLSPS 82  
 |||||  
 RESULT 13  
 AAB28398  
 ID AAB28398 standard; protein; 191 AA.  
 XX  
 AC AAB28398;  
 XX  
 DT 19-FEB-2001 (first entry)  
 XX  
 DE Murine endostatin.  
 XX  
 KW Murine; endostatin; cytostatic; antiproliferative;  
 KW vascular endothelial growth factor; VEGF; antibody; VEGF2 receptor;  
 KW cancer; vascularised solid tumour.  
 XX  
 OS Mus sp.  
 XX  
 FN WO2000064946-A2.  
 XX  
 PD 02-NOV-2000.  
 XX  
 PF 28-APR-2000; 2000WO-US011367.  
 XX  
 PR 28-APR-1999; 99US-0131432P.  
 XX  
 PA (TEXA) UNIV TEXAS SYSTEM.  
 XX  
 PI Thorpe PE, Brekken RA;  
 XX  
 DR WPI; 2000-687317/57.  
 DR N-PSDB; AAC67777.  
 XX  
 PT Immunogenic composition for the treatment and diagnosis of cancer  
 PT comprises an anti-VEGF (vascular endothelial growth factor) antibody  
 PT binding the same epitope as the monoclonal antibody ATCC PTA 1595.  
 XX  
 PS Example 10; Page 290-291; 298pp; English.  
 XX  
 CC The present invention relates to anti-Vascular Endothelial Growth Factor  
 CC (VEGF) antibodies that bind to the same epitope as the monoclonal  
 CC antibody ATCC PTA 1595 and which significantly inhibit VEGF binding to  
 CC the VEGF receptor VEGFR2, without inhibiting VEGF binding to the VEGF  
 CC receptor VEGFR1. The present sequence is murine endostatin. Endostatin  
 CC may be conjugated onto the anti-VEGF antibodies of the present invention.  
 CC The anti-VEGF antibodies of the present invention are useful for the  
 CC treatment and diagnosis of cancer, especially vascularised solid tumours  
 XX  
 SQ Sequence 191 AA;  
 Query Match 100.0%; Score 68; DB 3; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 0.00073;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VPVNLKDEVLSPS 14  
 Db 76 VPVNLKDEVLSPS 89  
 |||||



RESULT 14  
AAU77950  
ID AAU77950 standard; protein; 191 AA.  
XX  
XX AAU77950;  
AC  
DT 02-JUL-2002 (first entry)  
XX  
XX Amino acid sequence for mouse endostatin.  
DE  
XX  
XX Mouse; immunoconjugate; anti-vascular endothelial growth factor antibody;  
KW anti-VEGF antibody; monoclonal antibody 2C3 ATCC PTA 1595; VEGF receptor;  
KW VEGFR2; KDR/Flk-1; VEGFR1; Flt-1; angiogenesis; macular degeneration;  
KW ocular neovascular disease; cancer; vascularised solid tumour; AIDS;  
KW metastatic tumour; endothelial cell proliferation; inflammatory disorder;  
KW atherosclerosis; diabetic retinopathy; corneal graft rejection;  
KW acquired immune deficiency syndrome; infection; restenosis; fungal ulcer;  
KW sickle cell anaemia; endometriosis; endostatin.  
XX  
XX Mus sp.

OS  
XX AU200179401-A.  
PN  
XX 06-DEC-2001.  
PD  
XX 12-OCT-2001; 2001AU-00079401.  
XX  
XX 28-APR-2000; 2000AU-00048049.  
PR  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX  
XX Thorpe PE, Brekken RA;  
PI  
XX WPI; 2002-281368/33.  
DR  
XX N-PSDB; ABK47719.  
XX  
XX Immunoconjugate compositions for treating cancer by inhibiting  
PT angiogenesis and for delivering a diagnostic agent to tumor, comprises  
PT anti-vascular endothelial growth factor antibody attached to a biological  
PT agent.  
XX  
XX Example 10; Page 11-12 (Sequence listing); 299pp; English.

PS  
XX The present invention relates to antibody-based compositions comprising  
CC an immunoconjugate such as anti-vascular endothelial growth factor (VEGF)  
CC antibody (Ab) (or its antigen-binding fragment), attached to a biological  
CC agent, where the Ab binds to the same epitope as the monoclonal antibody  
CC (MAB) 2C3 ATCC PTA 1595, and significantly inhibits VEGF binding to the  
CC VEGF receptor VEGFR2 (KDR/Flk-1) without inhibiting VEGF binding to the  
CC VEGF receptor VEGFR1 (Flt-1). The compositions of the invention are  
CC useful in therapy, and diagnosis, for inhibiting angiogenesis in an  
CC animal having ocular neovascular disease or macular degeneration, and for  
CC delivering a biological agent to a vascularised tumour. The compositions  
CC can also be used for treating cancer and subjects at risk of developing,  
CC a vascularised solid tumour, a metastatic tumour or metastases from a  
CC primary tumour. The composition is useful for specifically inhibiting  
CC VEGF-induced endothelial cell proliferation, without significantly  
CC inhibiting VEGF-induced macrophage, osteoclast or chondroblast function.  
CC The compositions can be used for treating various diseases such as  
CC inflammatory disorders, atherosclerosis, diabetic retinopathy,  
CC restenosis, acquired immune deficiency syndrome (AIDS), blood borne  
CC tumours, corneal graft rejection, Crohn's disease, fungal ulcers,  
CC infections, sickle cell anaemia, and endometriosis. The present sequence  
CC represents mouse endostatin. Endostatin may be attached or functionally  
CC associated with anti-VEGF antibodies

XX  
SQ Sequence 191 AA;

Query Match 100.0%; Score 68; DB 5; Length 191;  
Best Local Similarity 100.0%; Pred. No. 0.00073;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFIVNLKDEVLSPS 14  
Db |||||

76 VFIVNLKDEVLSPS 89

RESULT 15  
AAE02031  
ID AAE02031 standard; protein; 207 AA.  
XX  
XX AAE02031;  
AC  
XX 31-JUL-2001 (first entry)  
DT

XX Murine endostatin fused to N-terminal secretion signal.  
DE  
XX Murine; endostatin; fusion protein; nucleotide-binding domain; NBD;  
KW ligand-binding domain; LBD; transcription regulating domain; TRD; cancer;  
KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;  
KW gene regulation; gene therapy; cell proliferative disorder; psoriasis;  
KW pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis.

XX Mus sp.

XX WO200130843-A1.

XX 03-MAY-2001.

XX 23-OCT-2000; 2000WO-EP010430.

XX 25-OCT-1999; 99US-00433042.

XX 02-JUN-2000; 2000US-00586625.

XX (NOVS ) NOVARTIS AG.

XX (SCRI ) SCRIPPS RES INST.

XX Barbas CF, Kadan M, Beerli R;

XX WPI; 2001-308618/32.

XX N-PSDB; AAD06108.

XX New fusion protein containing nucleotide-binding and ligand-binding  
PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated  
PT control of gene expression.

XX Example 19; Page 209; 218pp; English.

XX The invention relates to fusion protein comprising a nucleotide-binding  
CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor  
CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl  
CC zinc finger protein (ZFP), or a modular part of it, that interacts  
CC specifically with a contiguous sequence of at least 3 nucleotides. The  
CC fusion protein functions as a ligand-activated transcriptional regulator.  
CC The fusion protein and the nucleic acid encoding it, are used to regulate  
CC gene expression, particularly in gene therapy for treating malignant cell  
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell  
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,  
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The  
CC fusion protein and its DNA are also useful for treating diseases caused  
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion  
CC protein can be designed to target any selected gene (endogenous or  
CC exogenous), and can be made to have different selectivity or specificity  
CC for endogenous or exogenous ligands. The present sequence is murine  
CC endostatin fused to an N-terminal secretion signal. The corresponding  
CC cDNA sequence was used in the construction of left and shuttle plasmids  
CC containing regulatable transgene cassettes for evaluation of Cys2-His2  
CC zinc finger DNA binding domain (DBD)-Oestrogen receptor (ER) LBD  
CC regulators

XX  
SQ Sequence 207 AA;

Query Match 100.0%; Score 68; DB 4; Length 207;  
Best Local Similarity 100.0%; Pred. No. 0.0008;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VFIVNLKDEVLSPS 14  
| | | | | | | | | |  
Db 92 VFIVNLKDEVLSPS 105

Search completed: March 10, 2005, 11:06:52  
Job time : 72.1053 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:07:00 ; Search time 79.0263 Seconds  
(without alignments)  
54.260 Million cell updates/sec

Title: US-09-766-412-30  
Perfect score: 66  
Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.ppt.*
2:	/cgn2_6/ptodata/2/pubpaa/US07 PUBCOMB.ppt.*
3:	/cgn2_6/ptodata/2/pubpaa/US05 NEW PUB.ppt.*
4:	/cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.ppt.*
5:	/cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.ppt.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.ppt.*
7:	/cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.ppt.*
8:	/cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.ppt.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB.ppt.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB.ppt.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB.ppt.*
12:	/cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.ppt.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB.ppt.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB.ppt.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB.ppt.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D PUBCOMB.ppt.*
17:	/cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.ppt.*
18:	/cgn2_6/ptodata/2/pubpaa/US11 NEW PUB.ppt.*
19:	/cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.ppt.*
20:	/cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.ppt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	9 US-09-766-412-30	Sequence 30, Appl
2	66	100.0	20	9 US-09-873-676-115	Sequence 115, App
3	66	100.0	20	9 US-09-174-282-1	Sequence 1, Appli
4	66	100.0	20	9 US-09-405-493-1	Sequence 1, Appli
5	66	100.0	20	9 US-09-174-516-1	Sequence 1, Appli
6	66	100.0	20	9 US-09-154-302-1	Sequence 1, Appli
7	66	100.0	20	14 US-10-131-241-43	Sequence 43, Appl
8	66	100.0	20	14 US-10-232-316-1	Sequence 1, Appli
9	66	100.0	20	14 US-10-042-347-1	Sequence 1, Appli
10	66	100.0	20	15 US-10-351-284-1	Sequence 1, Appli
11	66	100.0	20	15 US-10-607-501-3	Sequence 3, Appli
12	66	100.0	20	16 US-10-358-311-1	Sequence 1, Appli
13	66	100.0	42	9 US-09-373-938-3	Sequence 3, Appli

14	66	100.0	184	14	US-10-131-241-46	Sequence 46, Appl
15	66	100.0	184	14	US-10-292-418-18	Sequence 18, Appl
16	66	100.0	184	17	US-10-920-820-13	Sequence 13, Appl
17	66	100.0	185	13	US-10-036-869-36	Sequence 36, Appl
18	66	100.0	191	9	US-09-998-831-13	Sequence 13, Appl
19	66	100.0	191	14	US-10-373-561-13	Sequence 13, Appl
20	66	100.0	207	9	US-09-373-938-2	Sequence 2, Appli
21	66	100.0	207	13	US-10-080-797-3	Sequence 3, Appli
22	66	100.0	207	14	US-10-422-934-71	Sequence 71, Appl
23	62	93.9	44	15	US-10-468-759-1	Sequence 1, Appli
24	62	93.9	171	9	US-09-978-531-3	Sequence 3, Appli
25	62	93.9	178	14	US-10-131-241-60	Sequence 60, Appl
26	62	93.9	178	14	US-10-042-347-5	Sequence 5, Appli
27	62	93.9	179	14	US-10-131-241-57	Sequence 57, Appl
28	62	93.9	180	14	US-10-131-241-47	Sequence 47, Appl
29	62	93.9	180	14	US-10-131-241-56	Sequence 56, Appl
30	62	93.9	181	14	US-10-131-241-55	Sequence 55, Appl
31	62	93.9	182	9	US-09-998-831-14	Sequence 14, Appl
32	62	93.9	182	14	US-10-131-241-54	Sequence 54, Appl
33	62	93.9	182	14	US-10-042-347-3	Sequence 3, Appli
34	62	93.9	182	14	US-10-373-561-14	Sequence 14, Appl
35	62	93.9	183	9	US-09-873-676-2	Sequence 2, Appli
36	62	93.9	183	9	US-09-978-531-1	Sequence 1, Appli
37	62	93.9	183	13	US-10-080-797-1	Sequence 1, Appli
38	62	93.9	183	14	US-10-131-241-52	Sequence 52, Appl
39	62	93.9	183	14	US-10-292-418-4	Sequence 4, Appli
40	62	93.9	183	15	US-10-135-872B-11	Sequence 11, Appl
41	62	93.9	183	15	US-10-607-501-2	Sequence 2, Appli
42	62	93.9	183	17	US-10-920-820-12	Sequence 12, Appl
43	62	93.9	184	15	US-10-210-172-162	Sequence 162, App
44	62	93.9	184	16	US-10-449-609-4	Sequence 4, Appli
45	62	93.9	184	17	US-10-920-820-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-09-766-412-30  
; Sequence 30, Application US/09766412  
; Patent No. US20020103129A1  
; GENERAL INFORMATION:  
; APPLICANT: GE, Ruowen et al.  
; TITLE OF INVENTION: SMALL PEPTIDES HAVING ANTI-ANGIOGENIC AND ENDOTHELIAL CELL INHIB  
; TITLE OF INVENTION: ACTIVITY  
; FILE REFERENCE: 1781-0215P  
; CURRENT APPLICATION NUMBER: US/09/766,412  
; CURRENT FILING DATE: 2001-01-11  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 13  
; TYPE: PPT  
; ORGANISM: Mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Endo-2  
US-09-766-412-30

Query Match 100.0%; Score 66; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 9.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 1 QPVLHLVALNTPL 13

RESULT 2  
US-09-873-676-115  
; Sequence 115, Application US/09873676  
; Patent No. US20020077289A1  
; GENERAL INFORMATION:

```
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-25933)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 115
; LENGTH: 20
; TYPE: PRT
; ORGANISM: mammalian
US-09-873-676-115

Query Match          100.0%; Score 66; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 3
US-09-174-282-1
; Sequence 1, Application US/09174282
; Publication No. US20020086352A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/174,282
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/740,168
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
```

```
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
US-09-174-282-1

Query Match          100.0%; Score 66; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 4
US-09-405-499-1
; Sequence 1, Application US/09405499
; Patent No. US20020123458A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Endostatin Protein and Fragments Thereof
; FILE REFERENCE: 05213-0640
; CURRENT APPLICATION NUMBER: US/09/405,499
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-405-499-1

Query Match          100.0%; Score 66; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 5
US-09-174-516-1
; Sequence 1, Application US/09174516A
; Patent No. US20020127595A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Methods of Detecting Endostatin Protein
; FILE REFERENCE: 05213-0227
; CURRENT APPLICATION NUMBER: US/09/174,516A
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-174-516-1

Query Match          100.0%; Score 66; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
   |||||
Db 7 QPVLHLVALNTPL 19

RESULT 6
US-09-154-302-1
; Sequence 1, Application US/09154302
```

Patent No. US20020155987A1  
GENERAL INFORMATION:  
APPLICANT: O'Reilly, Michael  
APPLICANT: Folkman, M. Judah  
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions  
TITLE OF INVENTION: and Methods  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones & Askew, LLP  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/154,302  
FILING DATE:  
CLASSIFICATION:  
PRIORITY INFORMATION DATA:  
APPLICATION NUMBER: US 08/740,168  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714  
REFERENCE/DOCKET NUMBER: 05213-0223  
TELEPHONE: 404-818-3700  
TELEFAX: 404-818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Murine  
TISSUE TYPE: Collagen  
US-09-154-302-1

Query Match 100.0%; Score 66; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
DB 7 QPVLHLVALNTPL 19

## RESULT 7

US-10-131-241-43  
Sequence 43, Application US/10131241  
Publication No. US20030012792A1  
GENERAL INFORMATION:  
APPLICANT: Holaday, John W.  
APPLICANT: Fortier, Anne H.  
TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation  
TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers  
FILE REFERENCE: 05213-0344 43170-271565  
CURRENT APPLICATION NUMBER: US/10/131,241  
CURRENT FILING DATE: 2002-07-22  
PRIOR APPLICATION NUMBER: US 09/413,049  
PRIOR FILING DATE: 1999-10-06  
PRIOR APPLICATION NUMBER: US 09/316,802  
PRIOR FILING DATE: 1999-05-21

PRIOR APPLICATION NUMBER: US 60/086,586  
PRIOR FILING DATE: 1998-05-22  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 43  
LENGTH: 20  
TYPE: PRT  
ORGANISM: Murinae sp.  
US-10-131-241-43

Query Match 100.0%; Score 66; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
DB 7 QPVLHLVALNTPL 19

## RESULT 8

US-10-232-316-1  
Sequence 1, Application US/10232316  
Publication No. US20030087393A1  
GENERAL INFORMATION:  
APPLICANT: O'Reilly, Michael  
Folkman, M. Judah  
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions  
and Methods  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jones & Askew, LLP  
STREET: 191 Peachtree, 37th Floor  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/232,316  
FILING DATE: 03-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/740,168A  
FILING DATE: 22-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren, William L.  
REGISTRATION NUMBER: 36,714  
REFERENCE/DOCKET NUMBER: 05213-0223  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-818-3700  
TELEFAX: 404-818-3799  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Murine  
TISSUE TYPE: Collagen  
US-10-232-316-1

Query Match 100.0%; Score 66; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.00015;

```
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 9
US-10-042-347-1
; Sequence 1, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fr
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-042-347-1

Query Match 100.0%; Score 66; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 10
US-10-351-284-1
; Sequence 1, Application US/10351284
; Publication No. US20030219426A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions and Methods
; FILE REFERENCE: 05213-3051 (43170-282623)
; CURRENT APPLICATION NUMBER: US/10/351,284
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.

Query Match 100.0%; Score 66; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 11
US-10-607-501-3
; Sequence 3, Application US/10607501
; Publication No. US20040091465A1
; GENERAL INFORMATION:
; APPLICANT: Yim, Zachary
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions and Methods
; FILE REFERENCE: 05213-0791 (43170-286879)
; CURRENT APPLICATION NUMBER: US/10/607,501
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/391,630
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-607-501-3

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 12
US-10-358-311-1
; Sequence 1, Application US/10358311
; Publication No. US20040102372A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Methods of Detecting Endostatin Protein
; FILE REFERENCE: 05213-0320 (43170-282919)
; CURRENT APPLICATION NUMBER: US/10/358,311
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 09/174,516
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-358-311-1

Query Match 100.0%; Score 66; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
US-10-351-284-1

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 11
US-10-607-501-3
; Sequence 3, Application US/10607501
; Publication No. US20040091465A1
; GENERAL INFORMATION:
; APPLICANT: Yim, Zachary
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions and Methods
; FILE REFERENCE: 05213-0791 (43170-286879)
; CURRENT APPLICATION NUMBER: US/10/607,501
; CURRENT FILING DATE: 2003-06-26
; PRIOR APPLICATION NUMBER: US 60/391,630
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-607-501-3

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 12
US-10-358-311-1
; Sequence 1, Application US/10358311
; Publication No. US20040102372A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Methods of Detecting Endostatin Protein
; FILE REFERENCE: 05213-0320 (43170-282919)
; CURRENT APPLICATION NUMBER: US/10/358,311
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 09/174,516
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-358-311-1

Query Match 100.0%; Score 66; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 QPVLHLVALNTPL 13  
Db 7 QPVLHLVALNTPL 19

## RESULT 13

US-09-373-938-3  
; Sequence 3, Application US/09373938  
; Publication No. US20020115202A1  
; GENERAL INFORMATION:  
; APPLICANT: Hallenbeck, Paul  
; APPLICANT: Chen, Cheayun Theresa  
; TITLE OF INVENTION: ADENOVRAL VECTORS INCLUDING DNA SEQUENCES ENCODING ANGIOGENIC IN  
; FILE REFERENCE: 4-30899P1  
; CURRENT APPLICATION NUMBER: US/09/373,938  
; CURRENT FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 42  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-373-938-3

Query Match 100.0%; Score 66; DB 9; Length 42;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 30 QPVLHLVALNTPL 42

## RESULT 14

US-10-131-241-46  
; Sequence 46, Application US/10131241  
; Publication No. US20030012792A1  
; GENERAL INFORMATION:  
; APPLICANT: Holaday, John W.  
; APPLICANT: Fortier, Anne H.  
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer  
; FILE REFERENCE: 05213-0344 43170-271565  
; CURRENT APPLICATION NUMBER: US/10/131,241  
; CURRENT FILING DATE: 2002-07-22  
; PRIOR APPLICATION NUMBER: US 09/413,049  
; PRIOR FILING DATE: 1999-10-06  
; PRIOR APPLICATION NUMBER: US 09/316,802  
; PRIOR FILING DATE: 1999-05-21  
; PRIOR APPLICATION NUMBER: US 60/086,586  
; PRIOR FILING DATE: 1998-05-22  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Murinae sp.  
US-10-131-241-46

Query Match 100.0%; Score 66; DB 14; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 7 QPVLHLVALNTPL 19

## RESULT 15

US-10-292-418-18  
; Sequence 18, Application US/10292418  
; Publication No. US20030139365A1  
; GENERAL INFORMATION:

; APPLICANT: Lo, Kin-Ming  
; APPLICANT: Li, Yue  
; APPLICANT: Gillies, Stephen D  
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as  
; TITLE OF INVENTION: Immunofusins  
; FILE REFERENCE: LEX-006C1  
; CURRENT APPLICATION NUMBER: US/10/292,418  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/383,315  
; PRIOR FILING DATE: 1999-08-25  
; PRIOR APPLICATION NUMBER: US 60/097,883  
; PRIOR FILING DATE: 1998-08-25  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 184  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-292-418-18

Query Match 100.0%; Score 66; DB 14; Length 184;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
Db 7 QPVLHLVALNTPL 19

Search completed: March 10, 2005, 11:15:26  
Job time : 79.0263 secs

**This Page Blank (uspto)**



OC Mammalia; Eutheria;

```

OX NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC066080; AAH66080.1; -.
DR GO; GO:0005604; C:basement membrane; IDA.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0001525; P:angiogenesis; IMP.
DR GO; GO:0001886; P:endothelial cell morphogenesis; IDA.
DR GO; GO:0043065; P:positive regulation of apoptosis; IDA.
DR GO; GO:0030335; P:positive regulation of cell migration; IDA.
DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR008161; Clg helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; Cona-like lec_gl.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR003129; TSP N.
DR Pfam; PF01391; Collagen; 7.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD000007; Clg helix; 2.
DR SMART; SM00210; TSPN; 1.
DR Collagen.
RW
SQ SEQUENCE 1315 AA; 134245 MW; 15C862C1E385F03A CRC64;
Query Match 100.0%; Score 66; DB 2; Length 1315;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 1138 QPVLHLVALNTPL 1150
|||||
RESULT 3
Q6PIY4 PRELIMINARY; PRT; 1315 AA.
AC Q6PIY4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Procollagen, type XVIII, alpha 1.
GN Name=Coll18a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC064817; AAH64817.1; -.
DR GO; GO:0005604; C:basement membrane; IDA.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0001525; P:angiogenesis; IMP.
DR GO; GO:0001886; P:endothelial cell morphogenesis; IDA.
DR GO; GO:0043065; P:positive regulation of apoptosis; IDA.
DR GO; GO:0030335; P:positive regulation of cell migration; IDA.
DR GO; GO:0008284; P:positive regulation of cell proliferation; IDA.
DR InterPro; IPR008161; Clg helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008985; Cona-like lec_gl.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR003129; TSP N.
DR Pfam; PF01391; Collagen; 7.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD000007; Clg helix; 2.
DR SMART; SM00210; TSPN; 1.
DR Collagen.
RW
SQ SEQUENCE 1315 AA; 134203 MW; B8215602ACE7AD1F CRC64;
Query Match 100.0%; Score 66; DB 2; Length 1315;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 1138 QPVLHLVALNTPL 1150
|||||
RESULT 4
CAIH MOUSE
ID CAIH MOUSE STANDARD; PRT; 1774 AA.
AC F39061; Q60672; Q61437; Q62001; Q62002; Q9JG63;
DT 01-FEB-1995 (Rel. 31, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].
GN Name=Coll18a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC STRAIN=BALE/c; TISSUE=Liver;

```



DR InterPro; IPR000024; Fz domain.  
 DR InterPro; IPR001791; Laminin G.  
 DR InterPro; IPR003129; TSP\_N.  
 DR Pfam; PF01391; Collagen; 8.  
 DR Pfam; PF06121; DUF959; 1.  
 DR Pfam; PF06482; Endostatin; 1.  
 DR Pfam; PF01392; Fz; 1.  
 DR Pfam; PF02210; TSP\_N; 1.  
 DR ProDom; PD000007; C1g\_helix; 1.  
 DR SMART; SM00063; FRI; 1.  
 DR SMART; SM00282; LamG; 1.  
 DR SMART; SM00210; TSPN; 1.  
 DR PROSITE; PS00308; FZ; 1.  
 DR 3D-structure; Alternative promoter usage; Alternative splicing;  
 KW Cell adhesion; Collagen; Direct protein sequencing;  
 KW Extracellular matrix; Glycoprotein; Hydroxylation; Repeat; Signal;  
 KW Structural protein.  
 FT SIGNAL 1 26 Potential.  
 FT CHAIN 27 1774 Collagen alpha 1(XVIII) chain.  
 FT CHAIN 1591 1774 Endostatin.  
 FT CHAIN 245 433 TSP N-terminal.  
 FT DOMAIN 365 482 FZ.  
 FT DOMAIN 27 785 Nonhelical region 1 (NC1).  
 FT DOMAIN 786 812 Triple-helical region 1 (COL1).  
 FT DOMAIN 813 822 Nonhelical region 2 (NC2).  
 FT DOMAIN 823 896 Triple-helical region 2 (COL2).  
 FT DOMAIN 897 920 Nonhelical region 3 (NC3).  
 FT DOMAIN 921 1042 Triple-helical region 3 (COL3).  
 FT DOMAIN 1043 1065 Nonhelical region 4 (NC4).  
 FT DOMAIN 1066 1148 Triple-helical region 4 (COL4).  
 FT DOMAIN 1149 1162 Nonhelical region 5 (NC5).  
 FT DOMAIN 1163 1204 Triple-helical region 5 (COL5).  
 FT DOMAIN 1205 1217 Nonhelical region 6 (NC6).  
 FT DOMAIN 1218 1290 Triple-helical region 6 (COL6).  
 FT DOMAIN 1291 1300 Nonhelical region 7 (NC7).  
 FT DOMAIN 1301 1333 Triple-helical region 7 (COL7).  
 FT DOMAIN 1334 1345 Nonhelical region 8 (NC8).  
 FT DOMAIN 1346 1369 Triple-helical region 8 (COL8).  
 FT DOMAIN 1370 1376 Nonhelical region 9 (NC9).  
 FT DOMAIN 1377 1428 Triple-helical region 9 (COL9).

Query Match 100.0%; Score 66; DB 1; Length 1774;  
 Best Local Similarity 100.0%; Pred. No. 0.0085;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
 |||||  
 DB 1597 QPVLHLVALNTPL 1609

RESULT 5  
 Q8WX15 PRELIMINARY; PRT; 187 AA.  
 AC Q8WX15;  
 DT 01-WAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Collagen XVIII (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21409408; PubMed=11517600;  
 RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.;  
 RT "Inhibition effect in vitro of purified endostatin expressed in Pichia  
 pastoris."  
 RL Sheng Wu Gong Cheng Xue Bao 17:278-282 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Feng Y., Wu Y., Zhu X., Liu C.X., Ma Q.J.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF416592; AAL37720.1; -.  
 DR HSP; P39060; 1BNL.  
 DR Pfam; PF06482; Endostatin; 1.  
 KW Collagen.  
 DR NON\_TER 1  
 SQ SEQUENCE 187 AA; 20448 MW; 72B1047D85839CD3 CRC64;  
 Query Match 93.9%; Score 62; DB 2; Length 187;  
 Best Local Similarity 92.3%; Pred. No. 0.0043;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
 |||||  
 DB 11 QPVLHLVALNSPL 23

RESULT 6  
 Q8NG19 PRELIMINARY; PRT; 261 AA.  
 ID Q8NG19;  
 AC Q8NG19;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Multi-functional protein MFP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dou D.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF282883; AAMS2249.1; -.  
 DR HSP; P39060; 1BNL.  
 DR InterPro; IPR010515; Endostatin.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF06482; Endostatin; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 KW Kringle.  
 SQ SEQUENCE 261 AA; 26745 MW; CA60C920AF3E90E5 CRC64;  
 Query Match 93.9%; Score 62; DB 2; Length 261;  
 Best Local Similarity 92.3%; Pred. No. 0.0062;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
 |||||  
 DB 85 QPVLHLVALNSPL 97

RESULT 7  
 Q8N4S4 PRELIMINARY; PRT; 816 AA.  
 ID Q8N4S4;  
 AC Q8N4S4;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Similar to collagen, type XVIII, alpha 1 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC033715; AAH33715.1; -.  
 DR HSP; P39060; 1BNL.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:000817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.

DR InterPro; IPR010515; Endostatin.  
 DR Pfam; PF01391; Collagen; 5.  
 DR Pfam; PF06482; Endostatin; 1.  
 DR ProDom; PD000007; C1g\_helix; 2.  
 KW Collagen.  
 FT NON\_TER  
 SQ SEQUENCE 816 AA; 82553 MW; 5D539B2946694F86 CRC64;

Query Match 93.9%; Score 62; DB 2; Length 816;  
 Best Local Similarity 92.3%; Pred. No. 0.02;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13  
 |||||  
 Db 640 QPVLHLVALNSPL 652

## RESULT 8

Q6R241 PRELIMINARY; PRT; 1336 AA.  
 AC Q6R241;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Type XVIII collagen short variant.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;  
 RA Elamasa H., Snellman A., Rehn M., Autio-Harminen H., Pihlajaniemi T.;  
 RT "Characterization of the human type XVIII collagen gene and  
 RT proteolytic processing and tissue location of the variant containing a  
 RT frizzled motif.";  
 RL Matrix Biol. 22:427-442(2003).  
 DR EMBL; AY484967; AAR83298.1; JOINED.  
 DR EMBL; AY484969; AAR83298.1; JOINED.  
 DR EMBL; AY484970; AAR83298.1; JOINED.  
 DR EMBL; AY484971; AAR83298.1; JOINED.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR010515; Endostatin.  
 DR Pfam; PF01391; Collagen; 6.  
 DR ProDom; PD000007; TSP N.  
 DR SMART; SM00210; TSPN; 1.  
 KW Collagen.  
 SQ SEQUENCE 1336 AA; 135522 MW; 1287BFAB2D5034C3 CRC64;

Query Match 93.9%; Score 62; DB 2; Length 1336;  
 Best Local Similarity 92.3%; Pred. No. 0.034;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13  
 |||||  
 Db 1160 QPVLHLVALNSPL 1172

## RESULT 9

O93419 PRELIMINARY; PRT; 1344 AA.  
 AC O93419;  
 DT 01-NOV-1998 (TReMBLrel. 08, Created)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)

DE Collagen XVIII precursor.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98411346; PubMed=9738008; DOI=10.1074/jbc.273.39.25404;  
 RA Halfter W., Dong S., Schurer B., Cole G.J.;  
 RT "Collagen XVIII is a basement membrane heparan sulfate proteoglycan.";  
 RL J. Biol. Chem. 273:25404-25412(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Halfter W., Dong S.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF083440; AAC33294.2; --  
 DR HSRP; P39060; IENL.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; C1g\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR010515; Endostatin.  
 DR InterPro; IPR01791; Laminin G.  
 DR InterPro; IPR003129; TSP N.  
 DR Pfam; PF01391; Collagen; 8.  
 DR Pfam; PF06482; Endostatin; 1.  
 DR ProDom; PD000007; C1g\_helix; 3.  
 DR SMART; SM00282; LamG; 1.  
 DR SMART; SM00210; TSPN; 1.  
 KW Collagen; Signal.  
 FT SIGNAL 1 26 Potential.  
 SQ SEQUENCE 1344 AA; 137402 MW; 7AA366E4FE940CCD CRC64;

Query Match 93.9%; Score 62; DB 2; Length 1344;  
 Best Local Similarity 92.3%; Pred. No. 0.034;  
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13  
 |||||  
 Db 1167 QPALHLVALNTPL 1179

## RESULT 10

CAIH\_HUMAN STANDARD; PRT; 1516 AA.  
 ID CAIH\_HUMAN  
 AC P39060; Q9UK38; Q9Y6Q7; Q9Y6Q8;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 29-MAR-2004 (Rel. 43, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].  
 GN Name=COL18A1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT ILE-841.  
 RX MEDLINE=98164096; PubMed=9503365; DOI=10.1016/S0945-053X(98)90003-8;  
 RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;  
 RT "Complete primary structure of two variant forms of human type XVIII  
 RT collagen and tissue-specific differences in the expression of the  
 RT corresponding transcripts.";  
 RL Matrix Biol. 16:319-328(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,  
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,

RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,  
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,  
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,  
RA Minochima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
RA Schwarze M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,  
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,  
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
RA Lehrach H., Reinhardt R., Yaspo M.-L.,  
RT "The DNA sequence of human chromosome 21.";  
RT Nature 405:311-319 (2000).  
RN [3]  
RN SEQUENCE OF 834-1516 FROM N.A., AND VARIANT ARG-886.  
RP MEDLINE=94245237; PubMed=8188291;  
RX Oh S.P., Warman M.L., Seidin M.F., Cheng S., Knoll J.H., Timmons S.,  
RA Olsen B.R.;  
RA "Cloning of cDNA and genomic DNA encoding human type XVIII collagen  
RT and localization of the alpha 1(XVIII) collagen gene to mouse  
RT chromosome 10 and human chromosome 21.";  
RL Genomics 19:494-499 (1994).  
RN [4]  
RN SEQUENCE OF 1334-1516 FROM N.A.  
RP TISSUE=Placenta;  
RC Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.;  
RA "Cloning and expression of human endostatin gene in Escherichia  
RT coli.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RN INVOLVEMENT IN KNOBLOCH SYNDROME.  
RP MEDLINE=20400445; PubMed=10942434; DOI=10.1093/hmg/9.13.2051;  
RX Sertie A.L., Soesi V., Camargo A.A., Zatz M., Brahe C.,  
RA Passos-Bueno M.R.;  
RA "Collagen XVIII, containing an endogenous inhibitor of angiogenesis  
RT and tumor growth, plays a critical role in the maintenance of retinal  
RT structure and in neural tube closure.";  
RL Hum. Mol. Genet. 9:2051-2058 (2000).  
RN [6]  
RN VARIANTS ILE-841 AND ASN-1437.  
RP MEDLINE=21518361; PubMed=11606364;  
RX Iughetti P., Suzuki O., Godoi P.H., Alves V.A., Sertie A.L.,  
RA Zorick T., Soares F., Camargo A.A., Moreira E.S., di Loreto C.,  
RA Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.;  
RT "A polymorphism in endostatin, an angiogenesis inhibitor, predisposes  
RT for the development of prostatic adenocarcinoma.";  
RL Cancer Res. 61:7375-7378 (2001).  
RN [7]  
RN VARIANTS LEU-49; ARG-111; ILE-841 AND ARG-886, AND CHARACTERIZATION OF  
RP VARIANT ASN-1437.  
RX PubMed=14695535; DOI=10.1002/humu.10284;  
RA Menzel O., Bekkeheien R.C.J., Raymond A., Fukai N., Boye E.,  
RA Kostolanyi G., Aftinos S., Deutsch S., Scott H.S., Olsen B.R.,  
RA Antonarakis S.E., Guipponi M.;  
RT "Knobloch syndrome: novel mutations in COL18A1, evidence for genetic  
RT heterogeneity, and a functionally impaired polymorphism in  
RT endostatin.";  
RL Hum. Mutat. 23:77-84 (2004).  
CC -!- FUNCTION: COL18A1 probably plays a major role in determining the  
CC retinal structure as well as in the closure of the neural tube.  
CC -!- FUNCTION: Endostatin potentially inhibits endothelial cell  
CC proliferation and angiogenesis. May inhibit angiogenesis by  
CC binding to the heparan sulfate proteoglycans involved in growth  
CC factor signaling.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long; Synonyms=NC-493;  
CC IsoId=P39060-1; Sequence=Displayed;  
CC Name=Short; Synonyms=NC1-303;  
CC IsoId=P39060-2; Sequence=VSP\_001155; VSP\_001156;  
CC -!- TISSUE SPECIFICITY: Present in multiple organs with highest levels  
CC in liver, lung and kidney.  
CC -!- PTM: Prolines at the third position of the tripeptide repeating  
CC unit (G-X-Y) are hydroxylated in some or all of the chains.  
CC -!- POLYMORPHISM: There is an association between a polymorphism in

CC position 1437 and prostate cancer. Heterozygous Asn-1437  
CC individuals have a 2.5 times increased chance of developing  
CC prostate cancer as compared with homozygous Asp-1437 individuals.  
CC -!- DISEASE: Defects in COL18A1 are a cause of Knobloch syndrome (KNO)  
CC [MIM:267750]; an autosomal recessive disorder defined by the  
CC occurrence of high myopia, vitreoretinal degeneration with retinal  
CC detachment, macular abnormalities and occipital encephalocele.  
CC -!- SIMILARITY: Belongs to the fibril-associated collagens with  
CC interrupted helices (FACIT) family.  
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF018081; AAC39658.1; -;  
CC EMBL; AF018082; AAC39659.1; -;  
CC EMBL; AL163302; CAB90482.1; -;  
CC EMBL; L22548; AAA51864.1; -;  
CC EMBL; AF184060; AAF01310.1; ALT INIT.  
CC PDB; 1BNL; X-ray; A/B/C/D=1334-1511.  
CC GlycoSuiteDB; P39060; -;  
CC Genew; HGNC:2195; COL18A1.  
CC MIM; 120328; -;  
CC GO; GO:0005581; C:collagen; TAS.  
CC GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
CC GO; GO:0009887; P:organogenesis; TAS.  
CC GO; GO:0007601; P:visual perception; TAS.  
CC InterPro; IPR008161; C1g helix.  
CC InterPro; IPR008985; ConA like lec\_gl.  
CC InterPro; IPR001791; Laminin\_G.  
CC InterPro; IPR003129; TSPN.  
CC Pfam; PF01391; Collagen; 5.  
CC Pfam; PF06121; DUF959; 1.  
CC Pfam; PF06482; Endostatin; 1.  
CC Pfam; PF02210; TSP N; 1.  
CC ProDom; PD000007; C1g helix; 1.  
CC SMART; SM00210; TSPN; 1.  
CC 3D-structure; Alternative splicing; Cell adhesion; Collagen;  
CC Extracellular matrix; Glycoprotein; Hydroxylation; Polymorphism;  
CC Repeat; Signal; Structural protein.  
CC SIGNAL 1 23 Potential.  
CC CHAIN 24 1516 Collagen alpha 1(XVIII) chain.  
CC CHAIN 1334 1516 Endostatin.  
CC DOMAIN 221 409 TSP N-terminal.  
CC DOMAIN 410 516 Nonhelical region 1 (NC1).  
CC DOMAIN 517 550 Triple-helical region 1 (COL1).  
CC DOMAIN 551 560 Nonhelical region 2 (NC2).  
CC DOMAIN 561 640 Triple-helical region 2 (COL2).  
CC DOMAIN 641 664 Nonhelical region 3 (NC3).  
CC DOMAIN 665 786 Triple-helical region 3 (COL3).  
CC DOMAIN 787 809 Nonhelical region 4 (NC4).  
CC DOMAIN 810 892 Triple-helical region 4 (COL4).  
CC DOMAIN 893 906 Nonhelical region 5 (NC5).  
CC DOMAIN 907 948 Triple-helical region 5 (COL5).  
CC DOMAIN 949 961 Nonhelical region 6 (NC6).  
CC DOMAIN 962 1034 Triple-helical region 6 (COL6).  
CC DOMAIN 1035 1044 Nonhelical region 7 (NC7).  
CC DOMAIN 1045 1077 Triple-helical region 7 (COL7).  
CC DOMAIN 1078 1089 Nonhelical region 8 (NC8).  
CC DOMAIN 1090 1111 Triple-helical region 8 (COL8).  
CC DOMAIN 1112 1118 Nonhelical region 9 (NC9).  
CC DOMAIN 1119 1173 Triple-helical region 9 (COL9).  
CC DOMAIN 1174 1186 Nonhelical region 10 (NC10).  
CC DOMAIN 1187 1204 Triple-helical region 10 (COL10).  
CC DOMAIN 1205 1516 Nonhelical region 11 (NC11).  
CC CARBOHYD N-linked (GlcNAc...) (Potential). 68

```

FT CARBOHYD 129 129 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 164 164 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1329 1329 O-linked (GlcNAc...),
FT /FTid=VAR_000150.
FT By similarity.
FT DISULFID 1366 1506 By similarity.
FT DISULFID 1468 1498 Cell attachment site (Potential).
FT SITE 1095 1097 Missing (in isoform Short).
FT VARSPLIC 1 180 /FTid=VSP_001155.
FT VARSPLIC 181 215 HTTEAGTLRPTSPSPSLGRPWAPLPGSVPPSS -> MA
FT PCPFWRRRLDVLAPLLVLGLVRAASAP (in
FT isoform Short).
FT /FTid=VSP_001156.
FT Q -> L.
FT /FTid=VAR_018053.
FT G -> R.
FT /FTid=VAR_018054.
FT V -> I.
FT /FTid=VAR_018055.
FT P -> R.
FT /FTid=VAR_018056.
FT D -> N (decreased activity for binding
FT laminin; increased risk of developing
FT prostate cancer; in compound
FT heterozygotes may cause Knobloch syndrome
FT when in combination with a frameshift/
FT truncating mutation).
FT /FTid=VAR_012709.

Query Match 93.9%; Score 62; DB 1; Length 1516;
Best Local Similarity 92.3%; Pred. No. 0.039; 0; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 0;

QY 1 QPVLHLVALNTPL 13
DB 1340 QPVLHLVALNSPL 1352

RESULT 11
ID Q6RZ39 PRELIMINARY; PRT; 1516 AA.
AC Q6RZ39;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type XVII collagen middle variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;
RA Elanaa H., Snellman A., Rehn M., Autio-Harminen H., Pihlajaniemi T.;
RT "Characterization of the human type XVIII collagen gene and
RT proteolytic processing and tissue location of the variant containing a
RT frizzled motif.";
RL Matrix Biol. 22:427-442(2003).
DR EMBL; AY484968; AAR83296.1; JOINED.
DR EMBL; AY484969; AAR83296.1; JOINED.
DR EMBL; AY484970; AAR83296.1; JOINED.
DR EMBL; AY484971; AAR83296.1; -.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016020; Membrane; IEA.
DR GO; GO:0005198; F-structural molecule activity; IEA.
DR GO; GO:0004888; F-transmembrane receptor activity; IEA.
DR GO; GO:0007155; P-cell adhesion; IEA.
DR GO; GO:0007275; P-development; IEA.
DR GO; GO:0008817; P-phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008985; ConA like Lec_g1.
DR InterPro; IPR010363; DUF959.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF01391; Collagen; 6.
DR Pfam; PF06121; DUF959; 1.
DR Pfam; PF06482; Endostatin; 1.
DR Pfam; PF01392; Fz; 1.
DR ProDom; PD0000007; Clg_helix; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00308; FZ; 1.
KW Collagen.
SQ SEQUENCE 1516 AA; 177920 MW; 481F4F72176FA2A6 CRC64;

Query Match 93.9%; Score 62; DB 2; Length 1751;
Best Local Similarity 92.3%; Pred. No. 0.045;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 1575 QPVLHLVALNSPL 1587

```

```

DR Pfam; PF01391; Collagen; 6.
DR Pfam; PF06121; DUF959; 1.
DR Pfam; PF06482; Endostatin; 1.
DR ProDom; PD0000007; Clg_helix; 1.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1516 AA; 153779 MW; 38D0F23D3FD758A4 CRC64;

Query Match 93.9%; Score 62; DB 2; Length 1516;
Best Local Similarity 92.3%; Pred. No. 0.039; 0; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 0;

QY 1 QPVLHLVALNTPL 13
DB 1340 QPVLHLVALNSPL 1352

RESULT 12
ID Q6RZ40 PRELIMINARY; PRT; 1751 AA.
AC Q6RZ40;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Type XVIII collagen long variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22980396; PubMed=14614989; DOI=10.1016/S0945-053X(03)00073-8;
RA Elanaa H., Snellman A., Rehn M., Autio-Harminen H., Pihlajaniemi T.;
RT "Characterization of the human type XVIII collagen gene and
RT proteolytic processing and tissue location of the variant containing a
RT frizzled motif.";
RL Matrix Biol. 22:427-442(2003).
DR EMBL; AY484968; AAR83296.1; JOINED.
DR EMBL; AY484969; AAR83296.1; JOINED.
DR EMBL; AY484970; AAR83296.1; JOINED.
DR EMBL; AY484971; AAR83296.1; -.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016020; Membrane; IEA.
DR GO; GO:0005198; F-structural molecule activity; IEA.
DR GO; GO:0004888; F-transmembrane receptor activity; IEA.
DR GO; GO:0007155; P-cell adhesion; IEA.
DR GO; GO:0007275; P-development; IEA.
DR GO; GO:0008817; P-phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008985; ConA like Lec_g1.
DR InterPro; IPR010363; DUF959.
DR InterPro; IPR010515; Endostatin.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF01391; Collagen; 6.
DR Pfam; PF06121; DUF959; 1.
DR Pfam; PF06482; Endostatin; 1.
DR Pfam; PF01392; Fz; 1.
DR ProDom; PD0000007; Clg_helix; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00308; FZ; 1.
KW Collagen.
SQ SEQUENCE 1751 AA; 177920 MW; 481F4F72176FA2A6 CRC64;

Query Match 93.9%; Score 62; DB 2; Length 1751;
Best Local Similarity 92.3%; Pred. No. 0.045;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 1575 QPVLHLVALNSPL 1587

```

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";



RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC076669; AAH76669.1; -  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008161; Clg\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008985; ConA\_like\_lectin.  
 DR InterPro; IPR010515; Endostatin.  
 DR InterPro; IPR003129; TSP\_N.  
 DR Pfam; PF01391; Collagen\_7.  
 DR Pfam; PF06482; Endostatin\_1.  
 DR ProDom; PD000007; Clg\_helix; 3.  
 DR SMART; SM00210; TSPN; 1.  
 KW Collagen.  
 SQ SEQUENCE 1310 AA; 134223 MW; A369B287758B422 CRC64;

Query Match 92.4%; Score 61; DB 2; Length 1310;  
 Best Local Similarity 100.0%; Pred. No. 0.051;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 PVLHLVALNTPL 13  
 |||||  
 Db 1134 PVLHLVALNTPL 1145

Search completed: March 10, 2005, 11:09:30  
 Job time : 54.3421 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:00:49 ; Search time 11.2895 Seconds  
(without alignments)  
110.795 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66

Sequence: 1 QPVLHLVALNTP1 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	1315	2 A56101	collagen alpha 1(X)
2	66	100.0	1774	2 B56101	collagen alpha 1(X)
3	62	93.9	684	2 A53019	collagen alpha 1(X)
4	44	66.7	1388	2 A53317	collagen alpha 1(X)
5	43	65.2	1867	2 S22775	MOT1 protein - yea
6	40	60.6	176	2 C70602	hypothetical prote
7	40	60.6	342	2 AH1045	conserved hypothet
8	39	59.1	190	2 D75319	3-methyladenine gl
9	39	59.1	333	2 I40791	acetoaldehyde dehydrogen
10	39	59.1	425	2 AB0796	NADH2 dehydrogenas
11	39	59.1	533	2 G95541	probable cytochrom
12	38	57.6	103	2 T50954	hypothetical prote
13	38	57.6	160	2 E72631	hypothetical prote
14	38	57.6	258	2 T25958	hypothetical prote
15	38	57.6	336	2 T44958	hypothetical prote
16	38	57.6	617	2 T15104	hypothetical prote
17	38	57.6	2311	2 T06161	acetyl-CoA carboxy
18	38	57.6	3206	1 GNVSP7	genome polyprotein
19	37	56.1	102	2 S50397	probable membrane
20	37	56.1	191	1 Q0BE4	BHR1 protein - hu
21	37	56.1	519	2 A95542	probable cytochrom
22	37	56.1	728	1 TVHUSK	transforming prote
23	36	54.5	72	1 PNBPF6	P13 protein - phag
24	36	54.5	298	2 T22264	hypothetical prote
25	36	54.5	323	2 T48425	lipase-like protei
26	36	54.5	331	2 A83534	probable C4-dicarb
27	36	54.5	342	1 S56374	hypothetical 38.7K
28	36	54.5	342	2 E86110	hypothetical prote
29	36	54.5	342	2 G91269	hypothetical prote

## ALIGNMENTS

### RESULT 1

A56101 collagen alpha 1(XVIII) chain precursor, short splice form - mouse

N:Contains: endostatin

C:Species: Mus musculus (house mouse)

C>Date: 03-Oct-1995 #sequence revision 08-May-1998 #text\_change 09-Jul-2004

C:Accession: A56101; A58371; S72450; S65595

R:Rehn, M.; Pihlajaniemi, T.

J. Biol. Chem. 270, 4705-4711, 1995

A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tise

tif homologous to rat and Drosophila frizzled proteins.

A:Reference number: A56101; MUID:95181468; PMID:7876242

A:Accession: A56101

A:Molecule type: mRNA

A:Residues: 1-103 <REH1>

A:Cross-references: UNIPROT:P39061; GB:U14636; NID:G618427; PIDN:ACS2178.1; PID:G61842

R:Rehn, M.; Pihlajaniemi, T.

Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994

A:Title: Alphas(XVIII), a collagen chain with frequent interruptions in the collagenous

A:Reference number: A58371; MUID:94240112; PMID:8183894

A:Accession: A58371

A:Molecule type: mRNA

A:Residues: 1-928 <REH2>

A:Cross-references: GB:L16898; NID:G404754; PIDN:AAA37434.1; PID:G553894

R:Oh, S.P.; Kanagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.

submitted to the EMBL Data Library, August 1993

A:Reference number: S72450

A:Accession: S72450

A:Molecule type: mRNA

A:Residues: 28-687,'L',689-734,'F',736-751,'R',753-1315 <OHW>

A:Cross-references: EMBL:L22545; NID:G348968; PIDN:AAA19787.1; PID:G511298

R:Oh, S.P.; Kanagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.

Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994

A:Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xa

A:Reference number: A58370; MUID:94240111; PMID:8183893

A:Accession: S65595

A:Molecule type: mRNA

A:Residues: 28-1315 <OHS>

A:Cross-references: EMBL:L22545

C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit

lated and subsequently O-glycosylated.

C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per

C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of u

ay be useful in treating solid tumors.

C:Genetics:

A:Gene: MGI:Coll18a1

A:Cross-references: MGI:711175

A:Map position: 10:41.0

C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteogly;

F.1-25/Domain: signal sequence #status predicted <SIG>

F.24-235/Region: thrombospondin amino-terminal homologous

F.26-1315/Product: collagen alpha 1(XVIII) chain, short splice form #status predicted <

F:327-353/Domain: collagenous #status predicted <CO1>  
 F:364-437/Domain: collagenous #status predicted <CO2>  
 F:462-583/Domain: collagenous #status predicted <CO3>  
 F:607-689/Domain: collagenous #status predicted <CO4>  
 F:704-745/Domain: collagenous #status predicted <CO5>  
 F:759-831/Domain: collagenous #status predicted <CO6>  
 F:842-874/Domain: collagenous #status predicted <CO7>  
 F:887-910/Domain: collagenous #status predicted <CO8>  
 F:892-894/Region: cell attachment (R-G-D) motif  
 F:918-969/Domain: collagenous #status predicted <CO9>  
 F:983-1000/Domain: collagenous #status predicted <CO10>  
 F:1132-1315/Product: endostatin #status predicted <EST>  
 F:1139-1315/Region: multiplexin collagen carboxyl-terminal homologous  
 F:126-488/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:172-228/Disulfide bonds: #status predicted  
 F:240,245,1257/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F:451,454,594/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 66; DB 2; Length 1315;  
 Best Local Similarity 100.0%; Pred. No. 0.0006; Mismatches 0; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13  
 |||||  
 DB 1138 QPVLHLVALNTPL 1150

RESULT 2  
 B56101  
 collagen alpha 1(XVIII) chain precursor, long splice form - mouse  
 N;Contains: collagen alpha 1(XVIII) chain precursor, medium splice form; endostatin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 03-Oct-1995 #sequence revision 08-May-1998 #text change 09-Jul-2004  
 C;Accession: B56101; C56101; S72450; S65595; PN0675; A54072; A58816  
 R;Rehn, M.; Pihlajaniemi, T.  
 J. Biol. Chem. 270, 4705-4711, 1995  
 A;Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue homologous to rat and Drosophila frizzled proteins.  
 A;Reference number: A56101; MUID:95181468; PMID:7876242  
 A;Accession: B56101  
 A;Molecule type: mRNA  
 A;Residues: 1-562 <REH1>  
 A;Cross-references: UNIPROT:Q61434; GB:U11637; NID:G618429; PIDN:AACS2179.1; PID:G618430  
 A;Experimental source: splice form clone PE17.24  
 A;Accession: C56101  
 A;Molecule type: mRNA  
 A;Residues: 1-239,487-562 <REH2>  
 A;Cross-references: GB:U11637; NID:G618429  
 A;Experimental source: splice form clones PE8.1, PE19, PE15.2  
 R;Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.  
 submitted to the EMBL Data Library, August 1993  
 A;Reference number: S72450  
 A;Accession: S72450  
 A;Molecule type: mRNA  
 A;Residues: 487-1146, 'L', 1148-1193, 'F', 1195-1210, 'R', 1212-1512, 'L', 1514-1522, 'F', 1524-1612, 'L', 1614-1628, 'L', 1630-1638, 'L', 1640-1648, 'L', 1650-1658, 'L', 1660-1668, 'L', 1670-1678, 'L', 1680-1688, 'L', 1690-1698, 'L', 1700-1708, 'L', 1710-1718, 'L', 1720-1728, 'L', 1730-1738, 'L', 1740-1748, 'L', 1750-1758, 'L', 1760-1768, 'L', 1770-1778, 'L', 1780-1788, 'L', 1790-1798, 'L', 1800-1808, 'L', 1810-1818, 'L', 1820-1828, 'L', 1830-1838, 'L', 1840-1848, 'L', 1850-1858, 'L', 1860-1868, 'L', 1870-1878, 'L', 1880-1888, 'L', 1890-1898, 'L', 1900-1908, 'L', 1910-1918, 'L', 1920-1928, 'L', 1930-1938, 'L', 1940-1948, 'L', 1950-1958, 'L', 1960-1968, 'L', 1970-1978, 'L', 1980-1988, 'L', 1990-1998, 'L', 2000-2008, 'L', 2010-2018, 'L', 2020-2028, 'L', 2030-2038, 'L', 2040-2048, 'L', 2050-2058, 'L', 2060-2068, 'L', 2070-2078, 'L', 2080-2088, 'L', 2090-2098, 'L', 2100-2108, 'L', 2110-2118, 'L', 2120-2128, 'L', 2130-2138, 'L', 2140-2148, 'L', 2150-2158, 'L', 2160-2168, 'L', 2170-2178, 'L', 2180-2188, 'L', 2190-2198, 'L', 2200-2208, 'L', 2210-2218, 'L', 2220-2228, 'L', 2230-2238, 'L', 2240-2248, 'L', 2250-2258, 'L', 2260-2268, 'L', 2270-2278, 'L', 2280-2288, 'L', 2290-2298, 'L', 2300-2308, 'L', 2310-2318, 'L', 2320-2328, 'L', 2330-2338, 'L', 2340-2348, 'L', 2350-2358, 'L', 2360-2368, 'L', 2370-2378, 'L', 2380-2388, 'L', 2390-2398, 'L', 2400-2408, 'L', 2410-2418, 'L', 2420-2428, 'L', 2430-2438, 'L', 2440-2448, 'L', 2450-2458, 'L', 2460-2468, 'L', 2470-2478, 'L', 2480-2488, 'L', 2490-2498, 'L', 2500-2508, 'L', 2510-2518, 'L', 2520-2528, 'L', 2530-2538, 'L', 2540-2548, 'L', 2550-2558, 'L', 2560-2568, 'L', 2570-2578, 'L', 2580-2588, 'L', 2590-2598, 'L', 2600-2608, 'L', 2610-2618, 'L', 2620-2628, 'L', 2630-2638, 'L', 2640-2648, 'L', 2650-2658, 'L', 2660-2668, 'L', 2670-2678, 'L', 2680-2688, 'L', 2690-2698, 'L', 2700-2708, 'L', 2710-2718, 'L', 2720-2728, 'L', 2730-2738, 'L', 2740-2748, 'L', 2750-2758, 'L', 2760-2768, 'L', 2770-2778, 'L', 2780-2788, 'L', 2790-2798, 'L', 2800-2808, 'L', 2810-2818, 'L', 2820-2828, 'L', 2830-2838, 'L', 2840-2848, 'L', 2850-2858, 'L', 2860-2868, 'L', 2870-2878, 'L', 2880-2888, 'L', 2890-2898, 'L', 2900-2908, 'L', 2910-2918, 'L', 2920-2928, 'L', 2930-2938, 'L', 2940-2948, 'L', 2950-2958, 'L', 2960-2968, 'L', 2970-2978, 'L', 2980-2988, 'L', 2990-2998, 'L', 3000-3008, 'L', 3010-3018, 'L', 3020-3028, 'L', 3030-3038, 'L', 3040-3048, 'L', 3050-3058, 'L', 3060-3068, 'L', 3070-3078, 'L', 3080-3088, 'L', 3090-3098, 'L', 3100-3108, 'L', 3110-3118, 'L', 3120-3128, 'L', 3130-3138, 'L', 3140-3148, 'L', 3150-3158, 'L', 3160-3168, 'L', 3170-3178, 'L', 3180-3188, 'L', 3190-3198, 'L', 3200-3208, 'L', 3210-3218, 'L', 3220-3228, 'L', 3230-3238, 'L', 3240-3248, 'L', 3250-3258, 'L', 3260-3268, 'L', 3270-3278, 'L', 3280-3288, 'L', 3290-3298, 'L', 3300-3308, 'L', 3310-3318, 'L', 3320-3328, 'L', 3330-3338, 'L', 3340-3348, 'L', 3350-3358, 'L', 3360-3368, 'L', 3370-3378, 'L', 3380-3388, 'L', 3390-3398, 'L', 3400-3408, 'L', 3410-3418, 'L', 3420-3428, 'L', 3430-3438, 'L', 3440-3448, 'L', 3450-3458, 'L', 3460-3468, 'L', 3470-3478, 'L', 3480-3488, 'L', 3490-3498, 'L', 3500-3508, 'L', 3510-3518, 'L', 3520-3528, 'L', 3530-3538, 'L', 3540-3548, 'L', 3550-3558, 'L', 3560-3568, 'L', 3570-3578, 'L', 3580-3588, 'L', 3590-3598, 'L', 3600-3608, 'L', 3610-3618, 'L', 3620-3628, 'L', 3630-3638, 'L', 3640-3648, 'L', 3650-3658, 'L', 3660-3668, 'L', 3670-3678, 'L', 3680-3688, 'L', 3690-3698, 'L', 3700-3708, 'L', 3710-3718, 'L', 3720-3728, 'L', 3730-3738, 'L', 3740-3748, 'L', 3750-3758, 'L', 3760-3768, 'L', 3770-3778, 'L', 3780-3788, 'L', 3790-3798, 'L', 3800-3808, 'L', 3810-3818, 'L', 3820-3828, 'L', 3830-3838, 'L', 3840-3848, 'L', 3850-3858, 'L', 3860-3868, 'L', 3870-3878, 'L', 3880-3888, 'L', 3890-3898, 'L', 3900-3908, 'L', 3910-3918, 'L', 3920-3928, 'L', 3930-3938, 'L', 3940-3948, 'L', 3950-3958, 'L', 3960-3968, 'L', 3970-3978, 'L', 3980-3988, 'L', 3990-3998, 'L', 4000-4008, 'L', 4010-4018, 'L', 4020-4028, 'L', 4030-4038, 'L', 4040-4048, 'L', 4050-4058, 'L', 4060-4068, 'L', 4070-4078, 'L', 4080-4088, 'L', 4090-4098, 'L', 4100-4108, 'L', 4110-4118, 'L', 4120-4128, 'L', 4130-4138, 'L', 4140-4148, 'L', 4150-4158, 'L', 4160-4168, 'L', 4170-4178, 'L', 4180-4188, 'L', 4190-4198, 'L', 4200-4208, 'L', 4210-4218, 'L', 4220-4228, 'L', 4230-4238, 'L', 4240-4248, 'L', 4250-4258, 'L', 4260-4268, 'L', 4270-4278, 'L', 4280-4288, 'L', 4290-4298, 'L', 4300-4308, 'L', 4310-4318, 'L', 4320-4328, 'L', 4330-4338, 'L', 4340-4348, 'L', 4350-4358, 'L', 4360-4368, 'L', 4370-4378, 'L', 4380-4388, 'L', 4390-4398, 'L', 4400-4408, 'L', 4410-4418, 'L', 4420-4428, 'L', 4430-4438, 'L', 4440-4448, 'L', 4450-4458, 'L', 4460-4468, 'L', 4470-4478, 'L', 4480-4488, 'L', 4490-4498, 'L', 4500-4508, 'L', 4510-4518, 'L', 4520-4528, 'L', 4530-4538, 'L', 4540-4548, 'L', 4550-4558, 'L', 4560-4568, 'L', 4570-4578, 'L', 4580-4588, 'L', 4590-4598, 'L', 4600-4608, 'L', 4610-4618, 'L', 4620-4628, 'L', 4630-4638, 'L', 4640-4648, 'L', 4650-4658, 'L', 4660-4668, 'L', 4670-4678, 'L', 4680-4688, 'L', 4690-4698, 'L', 4700-4708, 'L', 4710-4718, 'L', 4720-4728, 'L', 4730-4738, 'L', 4740-4748, 'L', 4750-4758, 'L', 4760-4768, 'L', 4770-4778, 'L', 4780-4788, 'L', 4790-4798, 'L', 4800-4808, 'L', 4810-4818, 'L', 4820-4828, 'L', 4830-4838, 'L', 4840-4848, 'L', 4850-4858, 'L', 4860-4868, 'L', 4870-4878, 'L', 4880-4888, 'L', 4890-4898, 'L', 4900-4908, 'L', 4910-4918, 'L', 4920-4928, 'L', 4930-4938, 'L', 4940-4948, 'L', 4950-4958, 'L', 4960-4968, 'L', 4970-4978, 'L', 4980-4988, 'L', 4990-4998, 'L', 5000-5008, 'L', 5010-5018, 'L', 5020-5028, 'L', 5030-5038, 'L', 5040-5048, 'L', 5050-5058, 'L', 5060-5068, 'L', 5070-5078, 'L', 5080-5088, 'L', 5090-5098, 'L', 5100-5108, 'L', 5110-5118, 'L', 5120-5128, 'L', 5130-5138, 'L', 5140-5148, 'L', 5150-5158, 'L', 5160-5168, 'L', 5170-5178, 'L', 5180-5188, 'L', 5190-5198, 'L', 5200-5208, 'L', 5210-5218, 'L', 5220-5228, 'L', 5230-5238, 'L', 5240-5248, 'L', 5250-5258, 'L', 5260-5268, 'L', 5270-5278, 'L', 5280-5288, 'L', 5290-5298, 'L', 5300-5308, 'L', 5310-5318, 'L', 5320-5328, 'L', 5330-5338, 'L', 5340-5348, 'L', 5350-5358, 'L', 5360-5368, 'L', 5370-5378, 'L', 5380-5388, 'L', 5390-5398, 'L', 5400-5408, 'L', 5410-5418, 'L', 5420-5428, 'L', 5430-5438, 'L', 5440-5448, 'L', 5450-5458, 'L', 5460-5468, 'L', 5470-5478, 'L', 5480-5488, 'L', 5490-5498, 'L', 5500-5508, 'L', 5510-5518, 'L', 5520-5528, 'L', 5530-5538, 'L', 5540-5548, 'L', 5550-5558, 'L', 5560-5568, 'L', 5570-5578, 'L', 5580-5588, 'L', 5590-5598, 'L', 5600-5608, 'L', 5610-5618, 'L', 5620-5628, 'L', 5630-5638, 'L', 5640-5648, 'L', 5650-5658, 'L', 5660-5668, 'L', 5670-5678, 'L', 5680-5688, 'L', 5690-5698, 'L', 5700-5708, 'L', 5710-5718, 'L', 5720-5728, 'L', 5730-5738, 'L', 5740-5748, 'L', 5750-5758, 'L', 5760-5768, 'L', 5770-5778, 'L', 5780-5788, 'L', 5790-5798, 'L', 5800-5808, 'L', 5810-5818, 'L', 5820-5828, 'L', 5830-5838, 'L', 5840-5848, 'L', 5850-5858, 'L', 5860-5868, 'L', 5870-5878, 'L', 5880-5888, 'L', 5890-5898, 'L', 5900-5908, 'L', 5910-5918, 'L', 5920-5928, 'L', 5930-5938, 'L', 5940-5948, 'L', 5950-5958, 'L', 5960-5968, 'L', 5970-5978, 'L', 5980-5988, 'L', 5990-5998, 'L', 6000-6008, 'L', 6010-6018, 'L', 6020-6028, 'L', 6030-6038, 'L', 6040-6048, 'L', 6050-6058, 'L', 6060-6068, 'L', 6070-6078, 'L', 6080-6088, 'L', 6090-6098, 'L', 6100-6108, 'L', 6110-6118, 'L', 6120-6128, 'L', 6130-6138, 'L', 6140-6148, 'L', 6150-6158, 'L', 6160-6168, 'L', 6170-6178, 'L', 6180-6188, 'L', 6190-6198, 'L', 6200-6208, 'L', 6210-6218, 'L', 6220-6228, 'L', 6230-6238, 'L', 6240-6248, 'L', 6250-6258, 'L', 6260-6268, 'L', 6270-6278, 'L', 6280-6288, 'L', 6290-6298, 'L', 6300-6308, 'L', 6310-6318, 'L', 6320-6328, 'L', 6330-6338, 'L', 6340-6348, 'L', 6350-6358, 'L', 6360-6368, 'L', 6370-6378, 'L', 6380-6388, 'L', 6390-6398, 'L', 6400-6408, 'L', 6410-6418, 'L', 6420-6428, 'L', 6430-6438, 'L', 6440-6448, 'L', 6450-6458, 'L', 6460-6468, 'L', 6470-6478, 'L', 6480-6488, 'L', 6490-6498, 'L', 6500-6508, 'L', 6510-6518, 'L', 6520-6528, 'L', 6530-6538, 'L', 6540-6548, 'L', 6550-6558, 'L', 6560-6568, 'L', 6570-6578, 'L', 6580-6588, 'L', 6590-6598, 'L', 6600-6608, 'L', 6610-6618, 'L', 6620-6628, 'L', 6630-6638, 'L', 6640-6648, 'L', 6650-6658, 'L', 6660-6668, 'L', 6670-6678, 'L', 6680-6688, 'L', 6690-6698, 'L', 6700-6708, 'L', 6710-6718, 'L', 6720-6728, 'L', 6730-6738, 'L', 6740-6748, 'L', 6750-6758, 'L', 6760-6768, 'L', 6770-6778, 'L', 6780-6788, 'L', 6790-6798, 'L', 6800-6808, 'L', 6810-6818, 'L', 6820-6828, 'L', 6830-6838, 'L', 6840-6848, 'L', 6850-6858, 'L', 6860-6868, 'L', 6870-6878, 'L', 6880-6888, 'L', 6890-6898, 'L', 6900-6908, 'L', 6910-6918, 'L', 6920-6928, 'L', 6930-6938, 'L', 6940-6948, 'L', 6950-6958, 'L', 6960-6968, 'L', 6970-6978, 'L', 6980-6988, 'L', 6990-6998, 'L', 7000-7008, 'L', 7010-7018, 'L', 7020-7028, 'L', 7030-7038, 'L', 7040-7048, 'L', 7050-7058, 'L', 7060-7068, 'L', 7070-7078, 'L', 7080-7088, 'L', 7090-7098, 'L', 7100-7108, 'L', 7110-7118, 'L', 7120-7128, 'L', 7130-7138, 'L', 7140-7148, 'L', 7150-7158, 'L', 7160-7168, 'L', 7170-7178, 'L', 7180-7188, 'L', 7190-7198, 'L', 7200-7208, 'L', 7210-7218, 'L', 7220-7228, 'L', 7230-7238, 'L', 7240-7248, 'L', 7250-7258, 'L', 7260-7268, 'L', 7270-7278, 'L', 7280-7288, 'L', 7290-7298, 'L', 7300-7308, 'L', 7310-7318, 'L', 7320-7328, 'L', 7330-7338, 'L', 7340-7348, 'L', 7350-7358, 'L', 7360-7368, 'L', 7370-7378, 'L', 7380-7388, 'L', 7390-7398, 'L', 7400-7408, 'L', 7410-7418, 'L', 7420-7428, 'L', 7430-7438, 'L', 7440-7448, 'L', 7450-7458, 'L', 7460-7468, 'L', 7470-7478, 'L', 7480-7488, 'L', 7490-7498, 'L', 7500-7508, 'L', 7510-7518, 'L', 7520-7528, 'L', 7530-7538, 'L', 7540-7548, 'L', 7550-7558, 'L', 7560-7568, 'L', 7570-7578, 'L', 7580-7588, 'L', 7590-7598, 'L', 7600-7608, 'L', 7610-7618, 'L', 7620-7628, 'L', 7630-7638, 'L', 7640-7648, 'L', 7650-7658, 'L', 7660-7668, 'L', 7670-7678, 'L', 7680-7688, 'L', 7690-7698, 'L', 7700-7708, 'L', 7710-7718, 'L', 7720-7728, 'L', 7730-7738, 'L', 7740-7748, 'L', 7750-7758, 'L', 7760-7768, 'L', 7770-7778, 'L', 7780-7788, 'L', 7790-7798, 'L', 7800-7808, 'L', 7810-7818, 'L', 7820-7828, 'L', 7830-7838, 'L', 7840-7848, 'L', 7850-7858, 'L', 7860-7868, 'L', 7870-7878, 'L', 7880-7888, 'L', 7890-7898, 'L', 7900-7908, 'L', 7910-7918, 'L', 7920-7928, 'L', 7930-7938, 'L', 7940-7948, 'L', 7950-7958, 'L', 7960-7968, 'L', 7970-7978, 'L', 7980-7988, 'L', 7990-7998, 'L', 8000-8008, 'L', 8010-8018, 'L', 8020-8028, 'L', 8030-8038, 'L', 8040-8048, 'L', 8050-8058, 'L', 8060-8068, 'L', 8070-8078, 'L', 8080-8088, 'L', 8090-8098, 'L', 8100-8108, 'L', 8110-8118, 'L', 8120-8128, 'L', 8130-8138, 'L', 8140-8148, 'L', 8150-8158, 'L', 8160-8168, 'L', 8170-8178, 'L', 8180-8188, 'L', 8190-8198, 'L', 8200-8208, 'L', 8210-8218, 'L', 8220-8228, 'L', 8230-8238, 'L', 8240-8248, 'L', 8250-8258, 'L', 8260-8268, 'L', 8270-8278, 'L', 8280-8288, 'L', 8290-8298, 'L', 8300-8308, 'L', 8310-8318, 'L', 8320-8328, 'L', 8330-8338, 'L', 8340-8348, 'L', 8350-8358, 'L', 8360-8368, 'L', 8370-8378, 'L', 8380-8388, 'L', 8390-8398, 'L', 8400-8408, 'L', 8410-8418, 'L', 8420-8428, 'L', 8430-8438, 'L', 8440-8448, 'L', 8450-8458, 'L', 8460-8468, 'L', 8470-8478, 'L', 8480-8488, 'L', 8490-8498, 'L', 8500-8508, 'L', 8510-8518, 'L', 8520-8528, 'L', 8530-8538, 'L', 8540-8548, 'L', 8550-8558, 'L', 8560-8568, 'L', 8570-8578, 'L', 8580-8588, 'L', 8590-8598, 'L', 8600-8608, 'L', 8610-8618, 'L', 8620-8628, 'L', 8630-8638, 'L', 8640-8648, 'L', 8650-8658, 'L', 8660-8668, 'L', 8670-8678, 'L', 8680-8688, 'L', 8690-8698, 'L', 8700-8708, 'L', 8710-8718, 'L', 8720-8728, 'L', 8730-8738, 'L', 8740-8748, 'L', 8750-8758, 'L', 8760-8768, 'L', 8770-8778, 'L', 8780-8788, 'L', 8790-8798, 'L', 8800-8808, 'L', 8810-8818, 'L', 8820-8828, 'L', 8830-8838, 'L', 8840-8848, 'L', 8850-8858, 'L', 8860-8868, 'L', 8870-8878, 'L', 8880-8888, 'L', 8890-8898, 'L', 8900-8908, 'L', 8910-8918, 'L', 8920-8928, 'L', 8930-8938, 'L', 8940-8948, 'L', 8950-8958, 'L', 8960-8968, 'L', 8970-8978, 'L', 8980-8988, 'L', 8990-8998, 'L', 9000-9008, 'L', 9010-9018, 'L', 9020-9028, 'L', 9030-9038, 'L', 9040-9048, 'L', 9050-9058, 'L', 9060-9068, 'L', 9070-9078, 'L', 9080-9088, 'L', 9090-9098, 'L', 9100-9108, 'L', 9110-9118, 'L', 9120-9128, 'L', 9130-9138, 'L', 9140-9148, 'L', 9150-9158, 'L', 9160-9168, 'L', 9170-9178, 'L', 9180-9188, 'L', 9190-9198, 'L', 9200-9208, 'L', 9210-9218, 'L', 9220-9228, 'L', 9230-9238, 'L', 9240-9248, 'L', 9250-9258, 'L', 9260-9268, 'L', 9270-9278, 'L', 9280-9288, 'L', 9290-9298, 'L', 9300-9308, 'L', 9310-9318, 'L', 9320-9328, 'L', 9330-9338, 'L', 9340-9348, 'L', 9350-9358, 'L', 9360-9368, 'L', 9370-9378, 'L

A:Note: in the authors' translation, 482-Gly is not shown, residues 483-490 are shifted  
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 lated and subsequently O-glycosylated.  
 C:Comment: Different splice forms of collagen alpha 1(XVIII) may be involved in perivasc  
 C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un  
 ay be useful in treating solid tumors.

C:Genetics:  
 A:Gene: GDB:COL18A1  
 A:Cross-references: GDB:138752; OMIM:120328  
 A:Map position: 21q22.3-21q22.3  
 C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc  
 F:1-684/Product: collagen alpha 1(XVIII) chain (fragment) #status predicted <MAT>  
 F:1-59/Domain: collagenous (fragment) #status predicted <CO4>  
 F:74-115/Domain: collagenous #status predicted <CO5>  
 F:129-201/Domain: collagenous #status predicted <CO6>  
 F:212-244/Domain: collagenous #status predicted <CO7>  
 F:257-278/Domain: collagenous #status predicted <CO8>  
 F:262-264/Region: cell attachment (R-G-D) motif  
 F:286-340/Domain: collagenous #status predicted <CO9>  
 F:354-371/Domain: collagenous #status predicted <CO10>  
 F:502-684/Product: endostatin #status predicted <EST>  
 F:509-684/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 93.9%; Score 62; DB 2; Length 684;  
 Best Local Similarity 92.3%; Pred. NO. 0.0017;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13  
 Db 508 QPVLHLVALNSPL 520

## RESULT 4

A53317  
 Collagen alpha 1(XV) chain precursor - human  
 N:Alternate names: procollagen alpha 1(XV) chain  
 C:Species: Homo sapiens (man)  
 C:Date: 07-Jul-1995 #sequence revision 07-Jul-1995 #text\_change 09-Jul-2004  
 C:Accession: A53317; A53146; S28778  
 R:Kivirikko, S.; Heinaemeeki, P.; Rehn, M.; Honkanen, N.; Myers, J.C.; Pihlajaniemi, T.  
 J. Biol. Chem. 269, 4773-4779, 1994  
 A:Title: Primary structure of the alpha1 chain of human type XV collagen and exon-intron  
 A:Reference number: A53317; MUID:94148920; PMID:8106446  
 A:Accession: A53317  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1388 <KIV>  
 A:Cross-references: UNIPROT:P39059; GB:L25280  
 A:Note: nucleotide sequence and conceptual translation not complete  
 R:Muragaki, Y.; Abe, N.; Ninomiya, Y.; Olsen, B.R.; Ooshima, A.  
 J. Biol. Chem. 269, 4042-4046, 1994

A:Title: The human alpha1(XV) collagen chain contains a large amino-terminal non-triple  
 A:Reference number: A53146; MUID:94140817; PMID:8307960

A:Accession: A53146  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-9,'S',11-48,'V',50-94,'A',96-149,'A',151-203,'V',205-408,'A',410-569 <MUR>  
 A:Cross-references: GB:D1230; NID:9415605; PIDN:BA04762.1; PID:d1005294; PID:g460703  
 R:Myers, J.C.; Kivirikko, S.; Gordon, M.K.; Dion, A.S.; Pihlajaniemi, T.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 10144-10148, 1992  
 A:Title: Identification of a previously unknown human collagen chain, alpha1(XV), charac  
 A:Reference number: S28778; MUID:93066196; PMID:1279671

A:Accession: S28778  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 544-640,'P',642-811,'P',813-1252 <MYE>  
 C:Genetics:  
 A:Gene: GDB:COL15A1

A:Cross-references: GDB:132578; OMIM:120325  
 A:Map position: 9q21-q922  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-1388/Product: collagen alpha 1(XV) chain #status predicted <MAT>  
 F:1216-1388/Region: multiplexin collagen carboxyl-terminal homologous

Query Match 66.7%; Score 44; DB 2; Length 1388;  
 Best Local Similarity 66.7%; Pred. No. 8.7;  
 Matches 8; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 12  
 Db 1215 KPALHLAALNMP 1226

## RESULT 5

S22775  
 MOT1 protein - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein Lpf4c; protein YPL082c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
 C:Accession: S22775; S61106  
 R:Davis, J.D.; Kunisawa, R.; Thorne, J.  
 Mol. Cell. Biol. 12, 1879-1892, 1992  
 A:Title: A presumptive helicase (MOT1 gene product) affects gene expression and is requi  
 A:Reference number: S22775; MUID:92195335; PMID:1312673  
 A:Accession: S22775  
 A:Molecule type: DNA  
 A:Residues: 1-1867 <DAV>  
 A:Cross-references: UNIPROT:P32333; EMBL:M83224; NID:g171964; PIDN:AAA34786.1; PID:g1719  
 R:Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa  
 submitted to the EMBL Data Library, August 1995  
 A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.  
 A:Reference number: S59677  
 A:Accession: S61106  
 A:Molecule type: DNA  
 A:Residues: 1-1867 <HAL>  
 A:Cross-references: EMBL:U41849; NID:g1147608; PID:g1147612; MIPS:YPL082c  
 C:Genetics:  
 A:Gene: SGD:MOT1  
 A:Cross-references: SGD:S0006003; MIPS:YPL082c  
 A:Map position: 16L  
 C:Keywords: DNA binding; nucleus; transmembrane protein  
 F:700-716/Domain: transmembrane #status predicted <TM1>  
 F:1038-1054/Domain: transmembrane #status predicted <TM2>  
 F:1186-1202/Domain: transmembrane #status predicted <TM3>

Query Match 65.2%; Score 43; DB 2; Length 1867;  
 Best Local Similarity 69.2%; Pred. No. 18;  
 Matches 9; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 QPVLHLVALNTPL 13  
 Db 637 QPILHL--LNTPV 647

## RESULT 6

C70602  
 Hypothetical protein Rv1000 - Mycobacterium tuberculosis (strain H37Rv)  
 C:Species: Mycobacterium tuberculosis  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: C70602  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: C70602  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-176 <COL>  
 A:Cross-references: GB:Z94752; GB:AL123456; NID:g3261731; PIDN:CAB08143.1; PID:g2052135  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: Rv1000  
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1000

```

Query Match          60.6%; Score 40; DB 2; Length 176;
Best Local Similarity 63.6%; Pred. No. 5.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 12
   ||: ||: ||: ||:
Db 103 PVVHLTALGSP 113

RESULT 7
AH1045
conserved hypothetical protein yJek [imported] - Salmonella enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH1045
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
. S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH1045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06813.1; PID:gl6505463; GSPDB:GN00176
C:Genetics:
A:Gene: yJek
C:Superfamily: conserved hypothetical protein yodO

Query Match          60.6%; Score 40; DB 2; Length 342;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HLVALNTP 13
   ||: ||: ||: ||:
Db 3 HVLNTNTP 11

RESULT 8
D75319
3-methyladenine glycosidase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75319
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
. M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: AY5250; MUID:20036896; PMID:10567266
A:Accession: D75319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <WHI>
A:Cross-references: UNIPROT:Q9RSQ0; GB:AE002043; GB:AE000513; NID:g6459859; PIDN:AAF1162
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2074
A:Map position: 1
C:Superfamily: Bacillus subtilis DNA-3-methyladenine glycosidase homolog yxjJ

Query Match          59.1%; Score 39; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 13
   ||: ||: ||: ||:
Db 136 PELHLLAPETPL 147

```

```

RESULT 9
I40791
acetoin dehydrogenase (TPP-dependent) (EC 1.-.-.) beta chain - Clostridium magnum
C:Species: Clostridium magnum
C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 12-Jul-2004
C:Accession: I40791
R:Kruger, N.; Oppermann, P.B.; Lorenzl, H.; Steinbuechel, A.
J. Bacteriol. 176, 3614-3630, 1994
A>Title: Biochemical and molecular characterization of the Clostridium magnum acetoin de
A:Reference number: I40789; MUID:94266715; PMID:8206840
A:Accession: I40791
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-333 <RU>
A:Cross-references: UNIPROT:Q46143; GB:L31844; NID:g472324; PIDN:AAA21745.1; PID:g472327
C:Superfamily: pyruvate dehydrogenase, E1 component, beta subunit
C:Keywords: oxidoreductase

Query Match          59.1%; Score 39; DB 2; Length 333;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 13
   ||: ||: ||: ||:
Db 297 PVVRIGALNTP 308

RESULT 10
AB0796
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) - Salmonella enterica subsp. enterica sero
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0796
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
. S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07548.1; PID:gl6503540; GSPDB:GN00176
C:Genetics:
A:Gene: STY2546
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C:Keywords: oxidoreductase

Query Match          59.1%; Score 39; DB 2; Length 425;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTP 13
   ||: ||: ||: ||:
Db 412 QPLISLVQLATPL 424

RESULT 11
G96541
probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G96541
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

```

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G96541  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-533 <STO>  
A:Cross-references: UNIPROT:Q9LPSS; GB:AE005173; NID:g11054632; PIDN:AAG27877.1; GSPDB:G  
C:Keywords: heme; iron; metalloprotein  
F:457/Binding site: heme iron (Cys) (axial ligand) #status predicted  
C:Genetics:  
A:Gene: F17J6.6  
A:Map position: 1  
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology  
C:Keywords: heme; iron; metalloprotein  
F:457/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 59.1%; Score 39; DB 2; Length 533;  
Best Local Similarity 58.3%; Pred. No. 27;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 13  
|:|:|:|:|:|:  
Db 77 PLHLRAFNPI 88

RESULT 12

T50954  
hypothetical protein B24P7.90 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 08-Sep-2000  
C:Accession: T50954  
R:Schulte, U.; Algn, V.; Hohetseel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, S.; submitted to the Protein Sequence Database, July 2000  
A:Reference number: 225286  
A:Accession: T50954  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-103 <SCH>  
A:Cross-references: EMBL:AL389890; GSPDB:GN001116; NCSP:B24P7.90  
A:Experimental source: BAC clone B24P7; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B24P7.90  
A:Map position: 6  
C:Superfamily: Neurospora crassa hypothetical protein B24P7.90

Query Match 57.6%; Score 38; DB 2; Length 103;  
Best Local Similarity 58.3%; Pred. No. 6.9;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTP 12  
|:|:|:|:|:|:  
Db 70 QPVLHLQLESP 81

RESULT 13

E72631  
hypothetical protein APE1508 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: E72631  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyrum*  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: E72631  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-160 <KAW>  
A:Cross-references: UNIPROT:Q9YEU1; DBJ:AF000061; NID:G5104821; PIDN:BAAG0507.1; PID:GDI  
A:Experimental source: strain K1  
C:Genetics:

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 10:59:54 ; Search time 44.2895 Seconds  
(without alignments)  
127.183 Million cell updates/sec

Title: US-09-766-412-29  
Perfect score: 59  
Sequence: 1 SPHNSYIVLPI 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	69.5	249	2 Q91D06	Q91D06 tt virus.
2	41	69.5	412	2 Q92VB2	Q92VB2 rhizobium m
3	41	69.5	414	2 Q82G9	Q82G9 brucella su
4	41	69.5	417	2 Q8YEF0	Q8YEF0 brucella me
5	40	67.8	135	2 Q6V9B7	Q6V9B7 penicillium
6	40	67.8	299	2 Q63MP4	Q63MP4 burkholderi
7	40	67.8	408	2 Q6LPY1	Q6LPY1 photobacter
8	40	67.8	413	2 Q9J894	Q9J894 spodoptera
9	40	67.8	425	1 Y175_PVRHO	O59545 pyrococcus
10	40	67.8	442	2 Q89RQ4	Q89RQ4 bradyrhizob
11	39	66.1	93	2 Q8ZWF9	Q8ZWF9 pyrobaculum
12	39	66.1	160	2 Q9CRT2	Q9CRT2 mus musculus
13	39	66.1	171	2 Q9WUW5	Q9WUW5 rattus norv
14	39	66.1	226	2 Q9QZD2	Q9QZD2 rattus norv
15	39	66.1	316	2 Q946U8	Q946U8 zea mays (m
16	39	66.1	410	2 Q92UI9	Q92UI9 rhizobium m
17	39	66.1	419	2 Q98C10	Q98C10 rhizobium l
18	39	66.1	428	2 Q8UD14	Q8UD14 agrobacteri
19	39	66.1	466	2 Q65EY8	Q65EY8 bacillus li
20	39	66.1	466	2 Q8DWE7	Q8DWE7 streptococ
21	39	66.1	472	2 Q837Y1	Q837Y1 enterococc
22	39	66.1	487	2 O15648	O15648 trypanosoma
23	39	66.1	545	2 Q74LN7	Q74LN7 lactobacill
24	39	66.1	1106	2 Q9GU55	Q9GU55 drosophila
25	39	66.1	1140	2 Q61434	Q61434 mus musculu
26	39	66.1	1315	2 Q6NZK9	Q6NZK9 mus musculu
27	39	66.1	1315	2 Q6P1Y4	Q6P1Y4 mus musculu
28	39	66.1	1774	1 CA1H_MOUSE	P39061 mus musculu
29	39	66.1	2171	2 Q9V841	Q9V841 drosophila
30	38.5	65.3	818	1 DAP2_YEAST	P18962 saccharomyc
31	38.5	65.3	818	2 Q66R87	Q66R87 saccharomyc

32 38 64.4 208 2 Q01494 Q01494 podospora a  
33 38 64.4 286 2 Q7REF5 Q7REF5 plasmodium  
34 38 64.4 400 2 Q63T55 Q63T55 burkholderi  
35 38 64.4 414 2 Q8XXJ8 Q8XXJ8 ralstonia s  
36 38 64.4 421 2 Q87YX2 Q87YX2 pseudomonas  
37 38 64.4 425 2 Q913K0 Q913K0 pseudomonas  
38 38 64.4 446 2 Q88F11 Q88F11 pseudomonas  
39 38 64.4 1243 2 Q87B08 Q87B08 xylella fas  
40 37 62.7 102 2 Q49991 Q49991 lycopersico  
41 37 62.7 179 2 Q9KDA4 Q9KDA4 bacillus ha  
42 37 62.7 205 2 Q75L46 Q75L46 oryza sativ  
43 37 62.7 206 2 Q7S937 Q7S937 neurospora  
44 37 62.7 239 2 Q6ZF18 Q6ZF18 oryza sativ  
45 37 62.7 259 2 Q730F0 Q730F0 bacillus ce

#### ALIGNMENTS

RESULT 1  
Q91D06 PRELIMINARY; PRT; 249 AA.  
AC Q91D06;  
AC Q91D06;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE ORF4.  
GN Name=ORF4;  
OS TT virus.  
OC Viruses; ssDNA viruses; Anellovirus.  
OX NCBI\_TaxID=68887;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21488921; PubMed=11601907; DOI=10.1006/viro.2001.1097;  
RA Okamoto H., Nishizawa T., Iakunashi M., Asabe S., Tsuda F.,  
RA Yoshikawa A.;  
RT "Heterogeneous distribution of TT virus of distinct genotypes in  
RT multiple tissues from infected humans.";  
RL Virology 288:358-368(2001).  
DR EMBL; AB060592; BAB69897.1; -.  
DR Pfam; PF02957; TT\_ORF2; 1.  
SQ SEQUENCE 249 AA; 27721 MW; B9195F046567DCBB CRC64;

Query Match 69.5%; Score 41; DB 2; Length 249;  
Best Local Similarity 60.0%; Pred. No. 23;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLP 10  
||| |||:  
Db 179 SPKSYVIKP 188

RESULT:2  
Q92VB2 PRELIMINARY; PRT; 412 AA.  
ID Q92VB2;  
AC Q92VB2;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Conserved hypothetical membrane protein, paralogue of Y20848.  
GN ORFNames=SMB21292;  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OG Plasmid pSymb.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoeiter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,  
RA Golding B., Puehler A.;



Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPHNSYI 7  
 |||||  
 Db 26 SPHNSYI 32

RESULT 6

Q63MP4 PRELIMINARY; PRT; 299 AA.

AC Q63MP4;  
 DT 25-OCT-2004 (TREMELrel. 28, Created)  
 DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)  
 DE LysR family regulatory protein.  
 GN ORFNames=BPS0606;  
 OS Burkholderia pseudomallei K96243.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Burkholderia.  
 OX NCBI\_TaxID=272560;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K96243;  
 RX PubMed=15377794;

RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,  
 RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,  
 RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,  
 RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,  
 RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshaizer D.,  
 RA Felwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagers K.,  
 RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,  
 RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,  
 RA Songvilai S., Stevens C., Tumapa S., Vesaratchavest M.,  
 RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.,  
 RT "Genomic plasticity of the causative agent of melioidosis,  
 RT Burkholderia pseudomallei";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).  
 DR EMBL; BX571966; CAH38063.1; -;  
 SQ SEQUENCE 299 AA; 32985 MW; 5993F6250CF3C589 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 299;  
 Best Local Similarity 63.6%; Pred. No. 42;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPHNSYI 11  
 :|||:|  
 Db 262 APHNYAVTPI 272

RESULT 7

Q6LPY1 PRELIMINARY; PRT; 408 AA.

AC Q6LPY1;  
 DT 05-JUL-2004 (TREMELrel. 27, Created)  
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)  
 DE Hypothetical protein ATU2314.  
 GN Name=ATU2314; OrderedLocusNames=PBPR2259;  
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Photobacterium.  
 OX NCBI\_TaxID=74109;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,  
 RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,  
 RA Valle G.;  
 RT "Genome analysis of Photobacterium profundum reveals the complexity of  
 RT high pressure adaptations";  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR378670; CAG20645.1; -;  
 DR InterPro; IPR000345; Cytochrome\_c.  
 DR InterPro; IPR009056; Cytochrome\_c.

DR InterPro; IPR010389; DUF989.  
 DR InterPro; IPR006162; Ppantne\_S.  
 DR Pfam; PF06181; DUF989; 1.  
 DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN\_1.  
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 408 AA; 44670 MW; 2DBBC0D825F26CFE CRC64;

Query Match 67.8%; Score 40; DB 2; Length 408;  
 Best Local Similarity 54.5%; Pred. No. 58;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYI 11  
 :|||:|  
 Db 221 SRHNNYLPV 231

RESULT 8

Q9J894 PRELIMINARY; PRT; 413 AA.

AC Q9J894;  
 DT 01-OCT-2000 (TREMELrel. 15, Created)  
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE ORF41 alkaline exonuclease.  
 OS Spodoptera exigua nucleopolyhedrovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_TaxID=10454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20036646; PubMed=10567663;  
 RA IJkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,  
 RA Goldbach R.W., Vlak J.M.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF169823; AAF33571.1; -;  
 DR GO; GO:0004527; F:exonuclease activity; IEA.  
 KW Exonuclease.  
 SQ SEQUENCE 413 AA; 48138 MW; 1B03329CD9FB34B6 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 413;  
 Best Local Similarity 66.7%; Pred. No. 59;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 HNSYIVLPI 11  
 :|||:|  
 Db 232 YNSYVLPV 240

RESULT 9

Y175\_PYRHO STANDARD; PRT; 425 AA.

AC O59545;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypothetical UPF0004 protein PH1875.  
 GN OrderedLocusNames=PH1875;  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- SIMILARITY: Belongs to the UPF0004 family.  
 CC -1- SIMILARITY: Contains 1 TRAM domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP000007; BAA30997.1; ALT INIT.  
 DR InterPro: IPR006638; ELP3/MiAB/NiB.  
 DR InterPro: IPR006466; MiAB-like B.  
 DR InterPro: IPR007197; Radical\_SAM.  
 DR InterPro: IPR002792; TRAM.  
 DR InterPro: IPR005839; UPF0004.  
 DR Pfam: PF04055; Radical\_SAM; 1.  
 DR Pfam: PF01938; TRAM; 1.  
 DR Pfam: PF00919; UPF0004; 1.  
 DR SMART: SM00729; ELP3; 1.  
 DR TIGRFAMs: TIGR01578; MiAB-like-B; 1.  
 DR TIGRFAMs: TIGR00089; UPF0004; 1.  
 DR PROSITE: PS50926; TRAM; 1.  
 DR PROSITE: PS01278; UPF0004; 1.  
 KW Complete proteome; Hypothetical protein.  
 FT DOMAIN 366 425 TRAM.  
 SQ SEQUENCE 425 AA; 48632 MW; 4375F6C16A19AE2D CRC64;  
 Query Match 67.8%; Score 40; DB 1; Length 425;  
 Best Local Similarity 63.6%; Pred. No. 60;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLP 11  
 Db 134 SPRNVYFILPI 144  
 RESULT 10  
 Q89RQ4 PRELIMINARY; PRT; 442 AA.  
 AC Q89RQ4;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE B12708 protein.  
 GN OrderedLocusNames=b12708;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,  
 RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpo S., Teurloka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL: AP005944; BAC47973.1;  
 DR InterPro: IPR000345; CysC\_heme\_BS.

DR InterPro: IPR009056; Cytochrome\_c.  
 DR InterPro: IPR010389; DUF989.  
 DR Pfam: PF06181; DUF989; 1.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 442 AA; 48389 MW; 6882BFE598C907AE CRC64;  
 Query Match 67.8%; Score 40; DB 2; Length 442;  
 Best Local Similarity 54.5%; Pred. No. 63;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLP 11  
 Db 270 SVHNNYILPV 280  
 RESULT 11  
 Q82WF9 PRELIMINARY; PRT; 93 AA.  
 AC Q82WF9;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conserved within P. aerophilum.  
 GN OrderedLocusNames=PAE1807;  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum  
 aerophilum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).  
 DR EMBL: AE009841; AAL63743.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 93 AA; 9497 MW; 598864288C4EF8F3 CRC64;  
 Query Match 66.1%; Score 39; DB 2; Length 93;  
 Best Local Similarity 77.8%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLP 9  
 Db 45 SPHNSHITL 53  
 RESULT 12  
 Q9CRT2 PRELIMINARY; PRT; 160 AA.  
 AC Q9CRT2;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched  
 DE library, clone:320000M10 product:procollagen, type XVIII, alpha 1,  
 DE full insert sequence. (Fragment).  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).



Query Match 66.1%; Score 39; DB 2; Length 226;  
Best Local Similarity 81.8%; Pred. No. 49;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | |  
Db 206 SCHNSYIVLCI 216

## RESULT 15

Q946U8  
ID Q946U8 PRELIMINARY; PRT; 316 AA.  
AC Q946U8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Long cell-linked locus protein.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Graziano E.I., Stiefel V., Puigdomenech P.,  
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF374405; AAL28975.1; -.  
SQ SEQUENCE 316 AA; 34551 MW; 833126FE0DBE4B8B CRC64;

Query Match 66.1%; Score 39; DB 2; Length 316;  
Best Local Similarity 70.0%; Pred. No. 69;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLP 10  
| | | | | | | |  
Db 51 SPHSSYIYSP 60

Search completed: March 10, 2005, 11:09:28  
Job time : 46.2895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:00:49 ; Search time 9.55263 Seconds  
(without alignments)  
110.795 Million cell updates/sec

Title: US-09-766-412-29

Perfect score: 59

Sequence: 1 SPHNSYIVLPI 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	69.5	412	2	D95941
2	41	69.5	417	2	AF3448
3	40	67.8	432	1	F71200
4	39	66.1	209	1	HFF26
5	39	66.1	410	2	E95984
6	39	66.1	428	2	A12860
7	39	66.1	428	2	H97637
8	39	66.1	1315	2	A56101
9	39	66.1	1774	2	B56101
10	38.5	65.3	818	1	A30107
11	38	64.4	208	2	S5950
12	38	64.4	425	2	D83455
13	37	62.7	102	2	T07890
14	37	62.7	179	2	F83813
15	37	62.7	264	2	H70903
16	37	62.7	352	2	F83729
17	37	62.7	475	2	AC0898
18	37	62.7	519	2	T48498
19	37	62.7	528	2	T34417
20	37	62.7	634	2	I40217
21	37	62.7	985	2	T41135
22	37	62.7	1264	2	T19545
23	37	62.7	2894	2	C64474
24	36	61.0	208	2	G96758
25	36	61.0	358	2	A96164
26	36	61.0	358	2	AF3123
27	36	61.0	461	2	T32406
28	36	61.0	880	2	G90163
29	36	61.0	1663	1	C3RT

30	35	59.3	53	2	T11383
31	35	59.3	109	2	F84012
32	35	59.3	165	2	T24470
33	35	59.3	210	2	K87256
34	35	59.3	319	2	S44235
35	35	59.3	381	2	T20622
36	35	59.3	416	2	E64140
37	35	59.3	493	2	T04767
38	35	59.3	625	2	G70150
39	35	59.3	652	2	T39409
40	35	59.3	2076	2	T28915
41	34.5	58.5	388	2	S49445
42	34	57.6	131	2	E91276
43	34	57.6	149	2	T38058
44	34	57.6	158	2	A70738
45	34	57.6	212	2	H97826

## ALIGNMENTS

### RESULT 1

D95941

conserved hypothetical membrane protein, paralogue of Y20848 Smb21292 [imported] - Sino.

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C:Accession: D95941

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Herna:

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A>Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: D95941

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <KUT>

A:Cross-references: UNIPROT:Q92VB2; GB:AL591985; PIDN:CAC49196.1; PID:g15140681; GSPDB:

A:Experimental source: strain 1021, megaplasmid pSymb

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kias, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

C:Contents: annotation

C:Genetics:

A:Gene: Smb21292

A:Genome: plasmid

Query Match 69.5%; Score 41; DB 2; Length 412;

Best Local Similarity 54.5%; Pred. No. 5.9;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11

Db 225 STNNYLTLPV 235

### RESULT 2

AF3448

hypothetical membrane spanning protein BMEI1572 [imported] - Brucella melitensis (strain

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004

R:Delvecchio, V.G.; Kaparal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A>Title: The genome sequence of the facultative intracellular pathogen Brucella meliten

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3448

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-417 <KUR>  
A;Cross-references: UNIPROT:Q8YFF0; GB:AE008917; PIDN:AAL52753.1; PID:g17983585; GSPDB:G  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI1572  
A;Map position: 1

Query Match 69.5%; Score 41; DB 2; Length 417;  
Best Local Similarity 54.5%; Pred. No. 5.9;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
DB 225 STNNVLTLPV 235  
| | | | |

RESULT 3  
F71200  
hypothetical protein PH1875 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 12-Jul-2004  
C;Accession: F71200  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: F71200  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-432 <RAW>  
A;Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30997.1; PID:g3258314  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH1875

Query Match 67.8%; Score 40; DB 1; Length 432;  
Best Local Similarity 63.6%; Pred. No. 9.6;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
DB 141 SPRNVYFILPI 151  
| | | | |

RESULT 4  
HHPF26  
heat shock protein 26 - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 09-May-2004  
C;Accession: A02920  
R;Ingolia, T.D.; Craig, E.A.  
Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982  
A;Title: Four small Drosophila heat shock proteins are related to each other and to mam  
A;Reference number: A93909; MUID:82248004; PMID:6285380  
A;Accession: A02920  
A;Molecule type: DNA  
A;Residues: 1-209 <ING>  
A;Note: the authors translated the codon GAC for residue 15 as Glu, CAG for residue 22 a  
A;Comment: This small heat shock protein is related to alpha crystallin.  
C;Genetics:  
A;Gene: FlyBase:Hsp26  
A;Cross-references: FlyBase:FBgn0001225  
A;Map position: 3L (67B)  
C;Superfamily: alpha-crystallin-related small heat shock protein  
C;Keywords: heat shock; stress-induced protein

Query Match 66.1%; Score 39; DB 1; Length 209;  
Best Local Similarity 50.0%; Pred. No. 6.6;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PHSNSYIVLPI 11  
| | | | |  
DB 29 PHSRYVLLPL 38  
| | | | |

# RESULT 5 E95984

conserved hypothetical membrane protein, paralogue of Y21292 SMB20848 [imported] - Sino  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: E95984  
R;Finan, T.M.; Weidner, S.; Wong, K.; Rühmester, J.; Chain, P.; Vorholter, P.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo  
A;Reference number: A95842; MUID:21396508; PMID:11481431  
A;Accession: E95984  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-410 <KUR>  
A;Cross-references: UNIPROT:Q92UI9; GB:AL591985; PIDN:CAC49541.1; PID:g15141028; GSPDB:  
A;Experimental source: strain 1021, megaplasmid pSymb  
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: SMB20848  
A;Genome: plasmid

Query Match 66.1%; Score 39; DB 2; Length 410;  
Best Local Similarity 54.5%; Pred. No. 14;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
DB 225 SLHNNYLTLPV 235  
| | | | |

# RESULT 6 A12860

conserved hypothetical protein Atu2314 [imported] - Agrobacterium tumefaciens (strain C  
C;Species: Agrobacterium tumefaciens  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: A12860  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: A12860

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-428 <KUR>  
A;Cross-references: UNIPROT:Q8UD14; GB:AE008688; PIDN:AAL43303.1; PID:g17740794; GSPDB:  
A;Experimental source: strain C58 (Dupont)  
C;Genetics:  
A;Gene: Atu2314  
A;Map position: circular chromosome

Query Match 66.1%; Score 39; DB 2; Length 428;  
Best Local Similarity 54.5%; Pred. No. 15;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |





A;Residues: 487-1512,'L',1514-1522,'P',1524-1683,'V',1685-1774 <OH2>  
A;Cross-references: EMBL:L22545  
R;Abe, N.; Muragaki, Y.; Yoshitaka, H.; Inoue, H.; Ninomiya, Y.  
Biochem. Biophys. Res. Commun. 196, 576-582, 1993  
A;Title: Identification of a novel collagen chain represented by extensive interruptions  
A;Reference number: PN0675; MUID:94059075; PMID:8240330  
A;Accession: PN0675  
A;Molecule type: mRNA  
A;Residues: 635-1774 <ABE>  
R;Rehn, M.; Hintikka, E.; Pihlajaniemi, T.  
J. Biol. Chem. 269, 13929-13935, 1994  
A;Title: Primary structure of the alpha chain of mouse type XVIII collagen, partial str  
collagen chain.  
A;Reference number: A54072; MUID:94245707; PMID:8188673  
A;Accession: A54072  
A;Molecule type: DNA; mRNA  
A;Residues: 1293-1403,'R',1405-1774 <REH3>  
A;Cross-references: GB:U03714; NID:9487733; PIDN:AAA20657.1; PID:G487734  
R;O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W.S.; Flynn, E.; Bi  
Cell 88, 277-285, 1997  
A;Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.  
A;Reference number: A58816; MUID:97160848; PMID:9008168  
A;Accession: A58816  
A;Molecule type: protein  
A;Residues: 1591-1610 <ORE>  
A;Experimental source: hemangioendothelium cells  
A;Note: Inhibits endothelial cell proliferation  
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
lated and subsequently O-glycosylated.  
C;Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in per  
C;Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of ur  
ay be useful in treating solid tumors.  
C;Genetics:  
A;Gene: MGI:Col18a1  
A;Cross-references: MGI:711175  
A;Map position: 10:41.0  
A;Introns: 1295/3; 1310/1; 1331/1; 1345/3; 1386/3; 1437/1; 1461/3; 1505/3; 1516/3; 1599/  
A;Note: the list of introns is incomplete  
C;Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc  
F;1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, long splice form #status pred  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;786-812/Domain: collagenous #status predicted <CO01>  
F;823-896/Domain: collagenous #status predicted <CO02>  
F;921-1042/Domain: collagenous #status predicted <CO03>  
F;1066-1148/Domain: collagenous #status predicted <CO04>  
F;1163-1204/Domain: collagenous #status predicted <CO05>  
F;1218-1290/Domain: collagenous #status predicted <CO06>  
F;1301-1333/Domain: collagenous #status predicted <CO07>  
F;1346-1369/Domain: collagenous #status predicted <CO08>  
F;1351-1353/Region: cell attachment (R-G-D) motif  
F;1377-1428/Domain: collagenous #status predicted <CO09>  
F;1442-1459/Domain: collagenous #status predicted <CO10>  
F;1591-1774/Product: endostatin #status predicted <EST>  
F;1598-1774/Region: multiplexin collagen carboxyl-terminal homologous  
F;354,361,947/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;699,704,1716/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F;910,913,1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted  
Query Match 66.1%; Score 39; DB 2; Length 1774;  
Best Local Similarity 81.8%; Pred. No. 72;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SPHNSVIVLPI 11  
Db 1754 SCHNSVIVLPI 1764  
RESULT 10  
A30107  
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)  
N;Alternate names: protein YHR028c  
C;Species: Saccharomyces cerevisiae

C;Date: 07-Jun-1990 #sequence\_revision 30-May-1997 #text\_change 09-Jul-2004  
C;Accession: S46780; A30107  
R;Du, Z.  
submitted to the EMBL Data Library, June 1994  
A;Description: The sequence of *S. cerevisiae* cosmid 8082.  
A;Reference number: S46773  
A;Accession: S46780  
A;Molecule type: DNA  
A;Residues: 1-818 <DUZ>  
A;Cross-references: UNIPROT:P18962; EMBL:U10399; NID:G500689; PIDN:AAB68879.1; PID:G5006  
R;Roberts, C.J.; Pohl, G.; Rothman, J.H.; Stevens, T.H.  
J. Cell Biol. 108, 1363-1373, 1989  
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in  
A;Reference number: A30107; MUID:89174971; PMID:2647766  
A;Accession: A30107  
A;Molecule type: DNA  
A;Residues: 1-82,'H',84-124,'N',126-181,'LRRLET',189-199,'N',201-365,'DFKRGKERK',376-5/  
A;Cross-references: EMBL:X15484  
A;Note: The authors translated the codon ACC for residue 572 as Asn  
C;Genetics:  
A;Gene: SGD:DAP2; STE13; MIPS:YHR028c  
A;Cross-references: SGD:S0001070; MIPS:YHR028c  
A;Map position: 8R  
C;Superfamily: dipeptidyl-peptidase IV  
C;Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; transmembrane protein; yeast vac  
F;30-45/Domain: transmembrane #status predicted <TM>  
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 65.3%; Score 38.5; DB 1; Length 818;  
Best Local Similarity 63.6%; Pred. No. 38;  
Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
QY 2 PHNSYI-VLPI 11  
Db 398 PHNGYVDILPI 408  
RESULT 11  
SS9950  
het-c2 protein - Podospora anserina  
C;Species: Podospora anserina  
C;Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Jul-2004  
C;Accession: SS9950  
R;Saupe, S.; Descamps, C.; Turcq, B.; Begueret, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 5927-5931, 1994  
A;Title: Inactivation of the Podospora anserina vegetative incompatibility locus het-c,  
A;Reference number: SS9950; MUID:94286551; PMID:8016091  
A;Accession: SS9950  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-208 <SAU>  
A;Cross-references: UNIPROT:Q01494; EMBL:U05236; NID:G523337; PIDN:AAA20542.1; PID:G5233  
C;Genetics:  
A;Introns: 72/3; 112/2  
Query Match 64.4%; Score 38; DB 2; Length 208;  
Best Local Similarity 60.0%; Pred. No. 10;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PHNSYIVLPI 11  
Db 145 PHHSFLVKPI 154  
RESULT 12  
D83455  
hypothetical protein PA1513 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: D83455  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: D83455  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-425 <STO>  
 A:Cross-references: UNIPROT:Q9I3K0; GB:AE004580; NID:g9947468; PIDN:AAG0490  
 C:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA1513

Query Match 64.4%; Score 38; DB 2; Length 425;  
 Best Local Similarity 54.5%; Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RESULT 13  
 T07890 1 SPHNSYIVLPI 11  
 |||:||||  
 Db 221 SRHNNYFTLPV 231

Query Match 64.4%; Score 38; DB 2; Length 425;  
 Best Local Similarity 54.5%; Pred. No. 23;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RESULT 13  
 T07890 1 SPHNSYIVLPI 11  
 |||:||||  
 Db 221 SRHNNYFTLPV 231

RESULT 13  
 T07890 protein kinase (EC 2.7.1.1-) 5 - tomato (fragment)  
 C:Species: *Lycopersicon esculentum* (tomato)  
 C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 16-Aug-2004  
 C:Accession: T07890  
 R:Zhang, S.; Diener, T.O.  
 submitted to the EMBL Data Library, February 1997  
 A:Description: Isolation and characterization of seven serine/threonine protein kinase-coding clones from tomato  
 A:Reference number: Z16191  
 A:Accession: T07890  
 A:Molecule type: mRNA  
 A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Residues: 1-102 <ZHA>  
 A:Cross-references: UNIPROT:O49991; EMBL:U89682; NID:g2735253; PIDN:AAB93863.1; PID:g2735253  
 A:Experimental source: cultivar Rutgers; seedling  
 C:Genetics:  
 A:Gene: PK5  
 C:Superfamily: protein kinase homology  
 C:Keywords: phosphotransferase  
 Query Match 62.7%; Score 37; DB 2; Length 102;  
 Best Local Similarity 54.5%; Pred. No. 7.2;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RESULT 14  
 P83813 1 SPHNSYIVLPI 11  
 :|||:|  
 Db 45 TPRNSYCAMP 55

Query Match 62.7%; Score 37; DB 2; Length 102;  
 Best Local Similarity 54.5%; Pred. No. 7.2;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 RESULT 14  
 P83813 phosphatidyleserine synthase psaA [imported] - *Bacillus halodurans* (strain C-125)  
 C:Species: *Bacillus halodurans*  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: F83813  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiraoka, Y.  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its relationship to other members of the family Halobacteriaceae  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: F83813  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-179 <STO>  
 A:Cross-references: UNIPROT:Q9KDA4; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA050004  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: psaA

Query Match 62.7%; Score 37; DB 2; Length 179;  
 Best Local Similarity 77.8%; Pred. No. 13;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 RESULT 14  
 P83813 phosphatidyleserine synthase psaA [imported] - *Bacillus halodurans* (strain C-125)  
 C:Species: *Bacillus halodurans*  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: F83813  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiraoka, Y.  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its relationship to other members of the family Halobacteriaceae  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: F83813  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-179 <STO>  
 A:Cross-references: UNIPROT:Q9KDA4; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA050004  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: psaA

Query Match 62.7%; Score 37; DB 2; Length 179;  
 Best Local Similarity 77.8%; Pred. No. 13;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 RESULT 14  
 P83813 phosphatidyleserine synthase psaA [imported] - *Bacillus halodurans* (strain C-125)  
 C:Species: *Bacillus halodurans*  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C:Accession: F83813  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiraoka, Y.  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its relationship to other members of the family Halobacteriaceae  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: F83813  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-179 <STO>  
 A:Cross-references: UNIPROT:Q9KDA4; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA050004  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: psaA

QY 3 HNSYIVLPI 11  
 |||:||||  
 Db 129 HNSFIGLPI 137  
 RESULT 15  
 H70903 probable transcription regulator Rv0165c - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 C:Accession: H70903  
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: H70903  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-284 <COL>  
 A:Cross-references: GB:297050; GB:AL123456; NID:g3256008; PIDN:CAB09759.1; PID:e323021;  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv0165c

Query Match 62.7%; Score 37; DB 2; Length 264;  
 Best Local Similarity 50.0%; Pred. No. 21;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 PHNSYIVLPI 11  
 |||:||||  
 Db 117 PHRGHVLE 126

Search completed: March 10, 2005, 11:10:07  
 Job time : 10.5526 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:07:00 ; Search time 66.8684 Seconds  
(without alignments)  
54.260 Million cell updates/sec

Title: US-09-766-412-29  
Perfect score: 59  
Sequence: 1 SPHNSYIVLPI 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	59	100.0	11	US-09-766-412-29	Sequence 29, Appl
2	40	67.8	638	US-10-282-122A-48112	Sequence 48112, A
3	39	66.1	184	US-10-131-241-46	Sequence 46, Appl
4	39	66.1	184	US-10-292-418-18	Sequence 18, Appl
5	39	66.1	184	US-10-920-820-13	Sequence 13, Appl
6	39	66.1	191	US-09-998-831-13	Sequence 13, Appl
7	39	66.1	191	US-10-373-561-13	Sequence 13, Appl
8	39	66.1	201	US-10-128-714-3451	Sequence 3451, Ap
9	39	66.1	201	US-10-128-714-8451	Sequence 8451, Ap
10	39	66.1	207	US-09-373-938-2	Sequence 2, Appli
11	39	66.1	207	US-10-080-797-3	Sequence 3, Appli
12	39	66.1	207	US-10-422-934-71	Sequence 71, Appl
13	39	66.1	455	US-09-815-242-4956	Sequence 4956, Ap

14	39	66.1	466	15	US-10-282-122A-72515	Sequence 72515, A
15	39	66.1	472	9	US-09-815-242-10778	Sequence 10778, A
16	39	66.1	472	15	US-10-282-122A-42556	Sequence 42556, A
17	38.5	65.3	818	14	US-10-401-437-3	Sequence 3, Appli
18	38.5	65.3	818	14	US-10-402-312-3	Sequence 3, Appli
19	38.5	65.3	818	14	US-10-402-087-3	Sequence 3, Appli
20	38.5	65.3	818	14	US-10-401-436-3	Sequence 3, Appli
21	38.5	65.3	818	15	US-10-369-493-22029	Sequence 22029, A
22	38	64.4	81	15	US-10-767-701-38716	Sequence 38716, A
23	38	64.4	175	16	US-10-424-599-280514	Sequence 280514, A
24	38	64.4	199	15	US-10-424-599-255994	Sequence 255994, A
25	38	64.4	439	15	US-10-424-599-244694	Sequence 244694, A
26	38	64.4	628	15	US-10-282-122A-49283	Sequence 49283, A
27	37	62.7	59	15	US-10-424-599-263035	Sequence 263035, A
28	37	62.7	81	15	US-10-425-114-51493	Sequence 51493, A
29	37	62.7	81	16	US-10-767-701-58526	Sequence 58526, A
30	37	62.7	89	15	US-10-424-599-143691	Sequence 143691, A
31	37	62.7	163	16	US-10-437-963-114067	Sequence 114067, A
32	37	62.7	227	15	US-10-282-122A-61727	Sequence 61727, A
33	37	62.7	245	15	US-10-282-122A-61983	Sequence 61983, A
34	37	62.7	264	15	US-10-282-122A-64349	Sequence 64349, A
35	37	62.7	265	15	US-10-282-122A-62628	Sequence 62628, A
36	37	62.7	330	16	US-10-437-963-181865	Sequence 181865, A
37	37	62.7	377	15	US-10-240-145-166	Sequence 166, App
38	37	62.7	442	14	US-10-385-136-2	Sequence 2, Appli
39	37	62.7	442	14	US-10-385-136-4	Sequence 4, Appli
40	37	62.7	442	14	US-10-385-136-6	Sequence 6, Appli
41	37	62.7	442	14	US-10-385-136-8	Sequence 8, Appli
42	37	62.7	448	15	US-10-282-122A-73493	Sequence 73493, A
43	37	62.7	451	15	US-10-282-122A-59616	Sequence 59616, A
44	37	62.7	465	15	US-10-282-122A-52776	Sequence 52776, A
45	37	62.7	475	9	US-09-815-242-13774	Sequence 13774, A

## ALIGNMENTS

RESULT 1  
US-09-766-412-29  
; Sequence 29, Application US/09766412  
; Patent No. US20020103129A1  
; GENERAL INFORMATION:  
; APPLICANT: GE, Ruowen et al.  
; TITLE OF INVENTION: SMALL PEPTIDES HAVING ANTI-ANGIOGENIC AND ENDOTHELIAL CELL INHIB  
; FILE REFERENCE: 1781-0215P  
; CURRENT APPLICATION NUMBER: US/09/766,412  
; CURRENT FILING DATE: 2001-01-11  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Mammalian  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Endo-1  
US-09-766-412-29

Query Match 100.0%; Score 59; DB 9; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SPHNSYIVLPI 11  
Db 1 SPHNSYIVLPI 11  
RESULT 2  
US-10-282-122A-48112  
; Sequence 48112, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48112
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48112

Query Match          67.8%; Score 40; DB 15; Length 638;
Best Local Similarity 63.6%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 SPHNSYIVLPI 11
Db      151 TPHLSHVLPI 161

RESULT 3
US-10-131-241-46
; Sequence 46, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-46

Query Match          66.1%; Score 39; DB 14; Length 184;
Best Local Similarity 81.8%; Pred. No. 65;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SPHNSYIVLPI 11
Db      164 SCHNSYIVLCI 174

RESULT 4
US-10-292-418-18
; Sequence 18, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Lo, Kin-Ming
; APPLICANT: Li, Yue
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; TITLE OF INVENTION: Immunofusins
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,883
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-292-418-18

Query Match          66.1%; Score 39; DB 14; Length 184;
Best Local Similarity 81.8%; Pred. No. 65;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SPHNSYIVLPI 11
Db      164 SCHNSYIVLCI 174

RESULT 5
US-10-920-820-13
; Sequence 13, Application US/10920820
; Publication No. US20050009144A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer, Inc. and Pfizer Products, Inc.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING DISORDERS
; TITLE OF INVENTION: INVOLVING ANGIOGENESIS
; FILE REFERENCE: 3153.00234/PCI0790A
; CURRENT APPLICATION NUMBER: US/10/920,820
; CURRENT FILING DATE: 2004-08-17
; PRIOR APPLICATION NUMBER: US/09/938,391
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US PROV NO. 60/227,924
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 184
; TYPE: PRT
; ORGANISM: MURINE
US-10-920-820-13
```

Query Match 66.1%; Score 39; DB 17; Length 184;  
Best Local Similarity 81.8%; Pred. No. 65;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | | | |  
DB 164 SCHNSYIVLCI 174

RESULT 6  
US-09-998-831-13  
; Sequence 13, Application US/09998831  
; Patent No. US20020119153A1  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; TITLE OF INVENTION: INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998,831  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 09/561,108  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-998-831-13

Query Match 66.1%; Score 39; DB 9; Length 191;  
Best Local Similarity 81.8%; Pred. No. 68;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | | | |  
DB 171 SCHNSYIVLCI 181

RESULT 7  
US-10-373-561-13  
; Sequence 13, Application US/10373561  
; Publication No. US20030175276A1  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002582  
; CURRENT APPLICATION NUMBER: US/10/373,561  
; CURRENT FILING DATE: 2003-02-24  
; PRIOR APPLICATION NUMBER: US/09/561,499  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-10-373-561-13

Query Match 66.1%; Score 39; DB 14; Length 191;  
Best Local Similarity 81.8%; Pred. No. 68;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | | | | | | |

DB 171 SCHNSYIVLCI 181

RESULT 8  
US-10-128-714-3451  
; Sequence 3451, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Broshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3451  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-3451

Query Match 66.1%; Score 39; DB 14; Length 201;  
Best Local Similarity 60.0%; Pred. No. 72;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PHNSYIVLPI 11  
| | | | | | | | | |  
DB 138 PHHSEFVKPI 147

RESULT 9  
US-10-128-714-8451  
; Sequence 8451, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Broshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
US-10-128-714-8451

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8451
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8451

Query Match          66.1%; Score 39; DB 14; Length 201;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PHNSYIVLPI 11
Db 138 PHHSFVKPI 147

RESULT 10
US-09-373-938-2
; Sequence 2, Application US/09373938
; Publication No. US20020115202A1
; GENERAL INFORMATION:
; APPLICANT: Hallenbeck, Paul
; APPLICANT: Chen, Cheayun Theresa
; TITLE OF INVENTION: ADENOVRAL VECTORS INCLUDING DNA SEQUENCES ENCODING ANGIOGENIC IN
; FILE REFERENCE: 4-30899P1
; CURRENT APPLICATION NUMBER: US/09/373,938
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-373-938-2

Query Match          66.1%; Score 39; DB 9; Length 207;
Best Local Similarity 81.8%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11
Db 187 SCHNSYIVLCI 197

RESULT 11
US-10-080-797-3
; Sequence 3, Application US/10080797
; Publication No. US20020183253A1
; GENERAL INFORMATION:
; APPLICANT: Campochiaro, Peter A.
; APPLICANT: Dixon, Katharine H.
; APPLICANT: Brazzell, Remulus K.
; TITLE OF INVENTION: METHOD FOR TREATING OCULAR
; FILE REFERENCE: 4-31881A
; CURRENT APPLICATION NUMBER: US/10/080,797
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Mouse
US-10-080-797-3

Query Match          66.1%; Score 39; DB 13; Length 207;
Best Local Similarity 81.8%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11
Db 187 SCHNSYIVLCI 197

US-09-815-242-4956
; Sequence 4956, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4956
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-4956
```



Query Match 66.1%; Score 39; DB 9; Length 455;  
 Best Local Similarity 45.5%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLPI 11  
 DB 418 APHGGFLVLPV 428

RESULT 14  
 US-10-282-122A-72515  
 ; Sequence 72515, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 72515  
 ; LENGTH: 466  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus mutans  
 US-10-282-122A-72515

Query Match 66.1%; Score 39; DB 15; Length 466;  
 Best Local Similarity 45.5%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLPI 11  
 DB 421 APHGGFLVLPV 431

RESULT 15  
 US-09-815-242-10778  
 ; Sequence 10778, Application US/09815242  
 ; Patent No. US20020061569A1

GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10778  
 ; LENGTH: 472  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 US-09-815-242-10778  
 Query Match 66.1%; Score 39; DB 9; Length 472;  
 Best Local Similarity 45.5%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 SPHNSYIVLPI 11  
 DB 422 APHGGFLVLPV 432  
 Search completed: March 10, 2005, 11:15:26  
 Job time: 67.8684 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 10, 2005, 11:04:09 ; Search time 22.2895 Seconds  
(without alignments)  
36.840 Million cell updates/sec

Title: US-09-766-412-29  
Perfect score: 59  
Sequence: 1 SPHNSYIVLPI 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2.6/prodata/1/iaa/5A COMB.pap.\*  
2: /cgn2.6/prodata/1/iaa/5B COMB.pap.\*  
3: /cgn2.6/prodata/1/iaa/6A COMB.pap.\*  
4: /cgn2.6/prodata/1/iaa/6B COMB.pap.\*  
5: /cgn2.6/prodata/1/iaa/PCUS COMB.pap.\*  
6: /cgn2.6/prodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	11	3	US-09-385-442-29
2	39	66.1	140	4	US-09-270-767-48510
3	39	66.1	151	4	US-09-270-767-40030
4	39	66.1	151	4	US-09-270-767-55246
5	39	66.1	184	4	US-09-938-391A-13
6	39	66.1	191	3	US-09-561-500-13
7	39	66.1	191	3	US-09-561-108-13
8	39	66.1	191	4	US-09-561-526-13
9	39	66.1	191	4	US-09-561-499-13
10	39	66.1	191	4	US-09-998-831-13
11	39	66.1	191	4	US-09-561-005-13
12	39	66.1	195	1	US-08-159-784-2
13	39	66.1	231	4	US-09-134-000C-5509
14	38.5	65.3	818	4	US-09-462-845-3
15	38.5	65.3	818	4	US-10-402-312-3
16	38	64.4	441	4	US-09-252-991A-32209
17	37	62.7	346	4	US-09-540-236-3202
18	37	62.7	442	4	US-09-787-083-2
19	37	62.7	442	4	US-09-787-083-4
20	37	62.7	442	4	US-09-787-083-6
21	37	62.7	442	4	US-09-787-083-8
22	37	62.7	498	4	US-09-107-532A-5037
23	37	62.7	634	1	US-07-779-049-3
24	37	62.7	634	1	US-08-080-240-3
25	36	61.0	189	4	US-09-270-767-38313
26	36	61.0	189	4	US-09-270-767-53530
27	36	61.0	250	4	US-09-702-705-1677

28	36	61.0	250	4	US-09-736-457-1677	Sequence 1677, Ap
29	36	61.0	250	4	US-09-671-325-1677	Sequence 1677, Ap
30	36	61.0	250	4	US-09-658-824-1677	Sequence 1677, Ap
31	36	61.0	304	4	US-09-902-540-12683	Sequence 12683, A
32	36	61.0	710	4	US-09-248-796A-14136	Sequence 14136, A
33	36	61.0	877	2	US-08-916-917-2	Sequence 2, Appli
34	36	61.0	877	2	US-08-972-631-2	Sequence 2, Appli
35	36	61.0	877	2	US-08-972-629-2	Sequence 2, Appli
36	36	61.0	877	2	US-08-972-630-2	Sequence 2, Appli
37	36	61.0	877	2	US-08-672-211-2	Sequence 2, Appli
38	36	61.0	877	3	US-09-225-170-2	Sequence 2, Appli
39	35.5	60.2	254	3	US-09-199-637A-241	Sequence 241, App
40	35	59.3	60	4	US-09-621-976-5575	Sequence 5575, Ap
41	35	59.3	87	4	US-09-107-532A-5785	Sequence 5785, Ap
42	35	59.3	95	4	US-09-248-796A-24504	Sequence 24504, A
43	35	59.3	168	4	US-09-248-796A-23213	Sequence 23213, A
44	35	59.3	362	4	US-09-134-000C-5244	Sequence 5244, Ap
45	35	59.3	560	4	US-09-821-016-3	Sequence 3, Appli

## ALIGNMENTS

### RESULT 1

US-09-385-442-29  
; Sequence 29, Application US/09385442  
; Patent No. 6200954  
; GENERAL INFORMATION:  
; APPLICANT: Ge, Ruowen  
; APPLICANT: Kini, R. Manjunatha  
; TITLE OF INVENTION: Small Peptides Having Potent Anti-Angiogenic Activity  
; FILE REFERENCE: 1781-170P  
; CURRENT APPLICATION NUMBER: US/09/385,442  
; CURRENT FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 60/099,313  
; EARLIER FILING DATE: 1999-09-04  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: mammalian  
; FEATURE:  
; OTHER INFORMATION: Endo-1  
US-09-385-442-29

Query Match 100.0%; Score 59; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPHNSYIVLPI 11  
Db 1 SPHNSYIVLPI 11

### RESULT 2

US-09-270-767-48510  
; Sequence 48510, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48510  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-48510

```
Query Match      66.1%; Score 39; DB 4; Length 140;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      4 NSYIVLPI 11
DB      116 NSYIVLPI 123

RESULT 3
US-09-270-767-40030
; Sequence 40030, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40030
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40030

Query Match      66.1%; Score 39; DB 4; Length 151;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

QY      1 SPHNSYIVLP 10
DB      82 SPHNNHVNLP 91

RESULT 4
US-09-270-767-55246
; Sequence 55246, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55246
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-55246

Query Match      66.1%; Score 39; DB 4; Length 151;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

QY      1 SPHNSYIVLP 10
DB      82 SPHNNHVNLP 91

RESULT 5
US-09-938-391A-13
; Sequence 13, Application US/09938391A
; Patent No. 6803211
; GENERAL INFORMATION:
; APPLICANT: Pfizer, Inc. and Pfizer Products, Inc.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND TREATING DISORDERS
; INVOLVING ANGIOGENESIS
; FILE REFERENCE: 3153.00234/PC10790A
; CURRENT APPLICATION NUMBER: US/09/938,391A
```

```
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US PROV NO. 60/227,924
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 184
; TYPE: PRT
; ORGANISM: MURINE
US-09-938-391A-13

Query Match      66.1%; Score 39; DB 4; Length 184;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SPHNSYIVLPI 11
DB      164 SCHNSYIVLCI 174

RESULT 6
US-09-561-500-13
; Sequence 13, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-500-13

Query Match      66.1%; Score 39; DB 3; Length 191;
Best Local Similarity 81.8%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 SPHNSYIVLPI 11
DB      171 SCHNSYIVLCI 181

RESULT 7
US-09-561-108-13
; Sequence 13, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
```

## US-09-561-108-13

Query Match 66.1%; Score 39; DB 3; Length 191;  
Best Local Similarity 81.8%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
Db 171 SCHNSYIVLCI 181

## RESULT 8

US-09-561-526-13  
; Sequence 13, Application US/09561526  
; Patent No. 6416758  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002586  
; CURRENT APPLICATION NUMBER: US/09/561,526  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-526-13

Query Match 66.1%; Score 39; DB 4; Length 191;  
Best Local Similarity 81.8%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
Db 171 SCHNSYIVLCI 181

## RESULT 9

US-09-561-499-13  
; Sequence 13, Application US/09561499  
; Patent No. 6524583  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002582  
; CURRENT APPLICATION NUMBER: US/09/561,499  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-499-13

Query Match 66.1%; Score 39; DB 4; Length 191;  
Best Local Similarity 81.8%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
Db 171 SCHNSYIVLCI 181

## RESULT 10

US-09-998-831-13  
; Sequence 13, Application US/09998831  
; Patent No. 6676941  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY  
; TITLE OF INVENTION: INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/998,831  
; CURRENT FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: 09/561,108  
; PRIOR FILING DATE: 2000-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-998-831-13

Query Match 66.1%; Score 39; DB 4; Length 191;  
Best Local Similarity 81.8%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
Db 171 SCHNSYIVLCI 181

## RESULT 11

US-09-561-005-13  
; Sequence 13, Application US/09561005  
; Patent No. 6703020  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE METHODS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002585  
; CURRENT APPLICATION NUMBER: US/09/561,005  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/131,432  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC  
US-09-561-005-13

Query Match 66.1%; Score 39; DB 4; Length 191;  
Best Local Similarity 81.8%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
| | | | |  
Db 171 SCHNSYIVLCI 181

## RESULT 12

US-08-159-784-2  
; Sequence 2, Application US/08159784  
; Patent No. 5643783  
; GENERAL INFORMATION:  
; APPLICANT: Bjorn R. Olsen

;; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF  
 ;; NUMBER OF SEQUENCES: 9  
 ;; CORRESPONDENCE ADDRESS:  
 ;; ADDRESSEE: Fish & Richardson  
 ;; STREET: 225 Franklin Street  
 ;; CITY: Boston  
 ;; STATE: Massachusetts  
 ;; COUNTRY: U.S.A.  
 ;; ZIP: 02110-2804  
 ;; COMPUTER READABLE FORM:  
 ;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ;; COMPUTER: IBM PS/2 Model 50Z or 55SX  
 ;; OPERATING SYSTEM: MS-DOS (Version 5.0)  
 ;; SOFTWARE: WordPerfect (Version 5.1)  
 ;; CURRENT APPLICATION DATA:  
 ;; APPLICATION NUMBER: US/08/159,784  
 ;; FILING DATE: December 1, 1993  
 ;; CLASSIFICATION: 530  
 ;; PRIOR APPLICATION DATA:  
 ;; APPLICATION NUMBER:  
 ;; FILING DATE:  
 ;; ATTORNEY/AGENT INFORMATION:  
 ;; NAME: John F. Freeman  
 ;; REGISTRATION NUMBER: 29,066  
 ;; REFERENCE/DOCKET NUMBER: 00246/170001  
 ;; TELECOMMUNICATION INFORMATION:  
 ;; TELEPHONE: (617) 542-5070  
 ;; TELEFAX: (617) 542-8906  
 ;; TELEX: 200154  
 ;; INFORMATION FOR SEQ ID NO: 2:  
 ;; SEQUENCE CHARACTERISTICS:  
 ;; LENGTH: 195  
 ;; TYPE: amino acid  
 ;; STRANDEDNESS: N/A  
 ;; TOPOLOGY: N/A  
 ;; US-08-159-784-2

Query Match 66.1%; Score 39; DB 1; Length 195;  
 Best Local Similarity 81.8%; Pred. No. 17;  
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
 DB 175 SCHNSYIVLCI 185

RESULT 13  
 US-09-134-000C-5609  
 ; Sequence 5609, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6912  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5609  
 ; LENGTH: 231  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 ; US-09-134-000C-5609

Query Match 66.1%; Score 39; DB 4; Length 231;  
 Best Local Similarity 45.5%; Pred. No. 21;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPHNSYIVLPI 11  
 :|| :|||

Db 181 APHGGFLVLPV 191

RESULT 14  
 US-09-462-845-3  
 ; Sequence 3, Application US/09462845  
 ; Patent No. 6723550  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Estell, David A.  
 ; TITLE OF INVENTION: Proteases From Gram-Positive Organisms  
 ; FILE REFERENCE: GC382-US  
 ; CURRENT APPLICATION NUMBER: US/09/462,845  
 ; CURRENT FILING DATE: 2000-01-13  
 ; PRIOR APPLICATION NUMBER: PCT/US98/14647  
 ; PRIOR FILING DATE: 1998-07-14  
 ; PRIOR APPLICATION NUMBER: EP 97305237.7  
 ; PRIOR FILING DATE: 1998-07-15  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 818  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus subtilis  
 ; US-09-462-845-3

Query Match 65.3%; Score 38.5; DB 4; Length 818;  
 Best Local Similarity 63.6%; Pred. No. 98;  
 Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 2 PHNSYI-VLPI 11  
 ||||| :|||  
 Db 398 PHNGYVDILPI 408

RESULT 15  
 US-10-402-312-3  
 ; Sequence 3, Application US/10402312  
 ; Patent No. 6833261  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Estell, David A.  
 ; TITLE OF INVENTION: Proteases From Gram-Positive Organisms  
 ; FILE REFERENCE: GC382-US  
 ; CURRENT APPLICATION NUMBER: US/10/402,312  
 ; CURRENT FILING DATE: 2003-03-26  
 ; PRIOR APPLICATION NUMBER: PCT/US98/14647  
 ; PRIOR FILING DATE: 1998-07-14  
 ; PRIOR APPLICATION NUMBER: EP 97305237.7  
 ; PRIOR FILING DATE: 1998-07-15  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 818  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus subtilis  
 ; US-10-402-312-3

Query Match 65.3%; Score 38.5; DB 4; Length 818;  
 Best Local Similarity 63.6%; Pred. No. 98;  
 Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 2 PHNSYI-VLPI 11  
 ||||| :|||  
 Db 398 PHNGYVDILPI 408

Search completed: March 10, 2005, 11:11:29  
 Job time : 22.2895 secs